

# SEARCH REQUEST FORM

Requestor's Name: \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Date: \_\_\_\_\_ Phone: \_\_\_\_\_ Art Unit: \_\_\_\_\_

## Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

## STAFF USE ONLY

Date completed: 10-29-02  
Searcher: Beverly @ 4994  
Terminal time: 20  
Elapsed time: 0  
CPU time: \_\_\_\_\_  
Total time: 23  
Number of Searches: \_\_\_\_\_  
Number of Databases: 1

### Search Site

\_\_\_\_\_ STIC  
\_\_\_\_\_ CM-1  
\_\_\_\_\_ Pre-S

### Type of Search

\_\_\_\_\_ N.A. Sequence  
\_\_\_\_\_ A.A. Sequence  
\_\_\_\_\_ Structure  
\_\_\_\_\_ Bibliographic

### Vendors

\_\_\_\_\_ IG Suite  
\_\_\_\_\_ STN  
\_\_\_\_\_ Dialog  
\_\_\_\_\_ APS  
\_\_\_\_\_ Geninfo  
\_\_\_\_\_ SDC  
\_\_\_\_\_ DARC/Questel  
\_\_\_\_\_ Other CGN

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 16:41:53 : Search time 34 Seconds  
(without alignments)  
1659.573 Million cell updates/sec

Title: US-09-992-238-2

Perfect score: 2644  
Sequence: 1 MNTCTNSTRNSSHTCMP.....GTGEGTEGKIVPSYDSATFP 508

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq\_032802.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2641	99.9	508	AAU25605	Human G Protein-Co
2	2641	99.9	508	AA86428	Human brain SERALP
3	2641	99.9	508	AAU04369	Human G-protein co
4	2641	99.9	508	AA64126	Human G-protein co
5	2641	99.9	508	AAU04387	GPCR-Gs fusion pro
6	2637	99.7	508	AA61982	Human G-protein co
7	2636	99.7	508	AAU04384	Human G-protein co
8	791	29.9	204	AAU25561	Human G-protein-Co
9	391	14.8	407	AAU70343	Human G protein-Co
10	388.5	14.7	454	AAU58586	Human histamine H2
11	383.5	14.5	552	AA614847	Human novel G-prot

12	371	14.0	529	17	AA890989	Human adrenergic G
13	371	14.0	529	19	AAW70501	Human G-protein ad
14	371	14.0	529	21	AA57167	Human adrenergic r
15	340.5	12.9	422	22	AA870249	HTR1A protein. Ho
16	338.5	12.8	466	15	AA853073	Alpha 1c adrenergic
17	338.5	12.8	466	15	AA858680	Human alpha-1c adr
18	338.5	12.8	466	16	AA890038	Human alpha-1c adr
19	338.5	12.8	466	16	AA885944	Human alpha-1c adr
20	338.5	12.8	466	22	AAU05407	Human adrenoceptor
21	338.5	12.8	466	22	AA859664	Human alpha-1c adr
22	333.5	12.6	429	17	AAW03714	Human alpha-1C3 ad
23	333.5	12.6	466	15	AA858679	Human alpha-1C3 ad
24	333.5	12.6	466	16	AA890037	Human alpha-1C3 ad
25	333.5	12.6	466	16	AA873604	Human alpha-1C3 ad
26	333.5	12.6	466	16	AA873603	Human alpha-1C3 ad
27	333.5	12.6	466	21	AA57170	Sequence of HuAlph
28	333.5	12.6	499	17	AAW03713	Human alpha-1C2 ad
29	332	12.6	375	22	AAU29934	Novel human secret
30	325.5	12.3	466	22	AA859669	Bovine alpha-1b ad
31	324	12.3	421	22	AA856316	Non-endogenous hum
32	323.5	12.2	466	15	AA852832	Sequence of human
33	314.5	11.9	501	15	AA870996	Human alpha-1A adr
34	314.5	11.9	501	16	AA890040	Human alpha-1A adr
35	314.5	11.9	501	21	AA57168	Sequence of HuAlph
36	314.5	11.9	555	22	AAU05409	Human adrenoceptor
37	314.5	11.9	572	15	AA853071	Alpha 1a adrenergic
38	314.5	11.9	572	16	AA890039	Human alpha-1A adr
39	314.5	11.9	572	16	AA85945	Alpha-1A adrenergic
40	314.5	11.9	572	22	AA898563	Alpha 1 A/D adrena
41	314.5	11.9	572	22	AA859662	Human alpha-1A adr
42	311.5	11.8	572	15	AA858681	Human alpha-1A adr
43	306.5	11.6	560	22	AA859666	Rat alpha-1a adren
44	304.5	11.5	601	13	AA821931	D.melanogaster oct
45	304.5	11.5	601	22	AA863318	Drosophila melano

#### ALIGNMENTS

RESULT 1	AAU25605	standard; Protein: 508 AA.
ID	AAU25605	
XX	AAU25605	
AC	AAU25605	
XX	18-DEC-2001	(first entry)
XX	Human G Protein-Coupled Receptor (GPCR) polypeptide #52.	
DE	Human G Protein-Coupled Receptor (GPCR) polypeptide #52.	
XX	Human G-protein coupled receptor; GPCR; mental disorder; schizophrenia;	
KW	attention deficit disorder; anxiety; depression; bipolar disorder;	
KW	neurological disorder; Huntington's disease; dementia; obesity; anorexia;	
KW	metabolic disorder; Parkinson's disease; Tourette's syndrome; thrombosis;	
KW	type 2 diabetes; cardiovascular disorder; myocardial infarction; cancer;	
KW	cardiomyopathy; atherosclerosis; human immunodeficiency virus; HIV;	
KW	viral infection; immunostimulant; neuroleptic; nootropic; tranquiliser;	
KW	antidepressant; anorectic; gene therapy.	
XX		
OS	Homo sapiens.	
PN	WO200162797-A2.	
XX	30-AUG-2001.	
PD		
XX		
PF	23-FEB-2001; 2001WO-US05676.	
XX		
PR	23-FEB-2000; 2000US-0184247.	
PR	23-FEB-2000; 2000US-0184303.	
PR	23-FEB-2000; 2000US-0184304.	
PR	23-FEB-2000; 2000US-0184305.	
PR	23-FEB-2000; 2000US-0184397.	
PR	02-MAR-2000; 2000US-0186457.	
PR	03-MAR-2000; 2000US-0186810.	



QY 301 RGSEEVESSTVASDGSMEKKEGSKTEVENSMDKADKRTVEVNOCSIDLGEDDMFEGDDI 360  
DB 301 RGSEEVESSTVASDGSMEKKEGSKTEVENSMDKADKRTVEVNOCSIDLGEDDMFEGDDI 360  
QY 361 NFSEDDVEAVNIPESLPPSRNSNSNPDLPRCYOCKAKAVIIFIIISYVLSLGPYCLAV 420  
DB 361 NFSEDDVEAVNIPESLPPSRNSNSNPDLPRCYOCKAKAVIIFIIISYVLSLGPYCLAV 420  
QY 421 LAVWVDETVQPMVITIIIMLFLOCCIHPIYVGYGMHKTIKKEIDMLKKFCKEKRPK 480  
DB 421 LAVWVDETVQPMVITIIIMLFLOCCIHPIYVGYGMHKTIKKEIDMLKKFCKEKRPK 480  
QY 481 EDSDPDLPGTEGTEGKIVPSYDSATPP 508  
DB 481 EDSDPDLPGTEGTEGKIVPSYDSATPP 508

RESULT 3  
AAU04369  
ID AAU04369 standard; Protein; 508 AA.  
AC AAU04369;  
DT 23-OCT-2001 (first entry)  
DE Human G-protein coupled receptor, hRUP15.  
KW Human; G-protein coupled receptor; GPCR; hRUP15; agonist;  
KW Inverse agonist; lung cancer.  
XX Homo sapiens.  
OS  
PN MO200136471-A2.  
XX  
PD 25-MAY-2001.  
XX  
PF 16-NOV-2000; 2000MO-US31509.  
XX  
PR 17-NOV-1999; 99US-0166088.  
PR 17-NOV-1999; 99US-0166099.  
PR 17-NOV-1999; 99US-0166369.  
PR 23-DEC-1999; 99US-0171900.  
PR 23-DEC-1999; 99US-0171901.  
PR 23-DEC-1999; 99US-0171902.  
PR 11-FEB-2000; 2000US-0181749.  
PR 14-MAR-2000; 2000US-0189258.  
PR 14-MAR-2000; 2000US-0189259.  
PR 10-APR-2000; 2000US-0195898.  
PR 10-APR-2000; 2000US-0195899.  
PR 10-APR-2000; 2000US-0196078.  
PR 28-APR-2000; 2000US-0200419.  
PR 12-MAY-2000; 2000US-0203630.  
PR 12-JUN-2000; 2000US-0210741.  
PR 12-JUN-2000; 2000US-0210982.  
PR 21-AUG-2000; 2000US-0226760.  
PR 26-SEP-2000; 2000US-0235418.  
PR 26-SEP-2000; 2000US-0235779.  
PR 20-OCT-2000; 2000US-0242332.  
PR 20-OCT-2000; 2000US-0242343.  
XX  
PA (AREN-) ARENA PHARM INC.  
XX  
PI Chen R, Dang HT, Lowitz KP;  
XX  
DR MPI: 2001-355616/37.  
DR N-PSDB; AAS07942.  
XX  
PT Endogenous and non-endogenous versions of human G-protein coupled  
PT receptors for direct identification of candidate compounds as agonists,  
PT inverse agonists or partial agonists for use as therapeutic agents -  
XX  
PS Claim 29; Page 102-104; 160pp; English.

XX  
CC The sequence represents a human G-protein coupled receptor (GPCR),  
CC hRUP15. The endogenous and non-endogenous, constitutively activated  
CC versions of human G-protein coupled receptors (GPCR), are useful for  
CC direct identification of candidate compounds as receptor agonists,  
CC inverse agonists or partial agonists having applicability as therapeutic  
CC agents for treating diseases related to GPCR, e.g. lung cancer.  
CC Non-endogenous version of human GPCRs are also utilized in research  
CC settings and in vitro and in vivo system, incorporating GPCRs can be  
CC utilized to elucidate and understand the roles these receptors  
CC play in the human condition, both normal and diseased.  
XX  
QY Sequence 508 AA:  
Query Match 99.9%; Score 2641; DB 22; Length 508;  
Best Local Similarity 99.8%; Pred. No. 6, 5e-234;  
Matches 507; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MTSTCTNSTRESNSHRCMPLSKPIISLAGIIRSTVLFVFLAASFVGNIVLVLQRP 60  
DB 1 MTSTCTNSTRESNSHRCMPLSKPIISLAGIIRSTVLFVFLAASFVGNIVLVLQRP 60  
QY 61 QLLQVTRNFIFNLVTDLQISLVAPWVATSVPLFPLNSHFCALVSLTLFAFASVN 120  
DB 61 QLLQVTRNFIFNLVTDLQISLVAPWVATSVPLFPLNSHFCALVSLTLFAFASVN 120  
QY 121 TIVVSVDRILSTIHPISTYSKMTQRRGYLLYGTWVAILQSTIPPLYGNGQAAFDERNA 180  
DB 121 TIVVSVDRILSTIHPISTYSKMTQRRGYLLYGTWVAILQSTIPPLYGNGQAAFDERNA 180  
QY 181 LCSMIMGASPSYITLISVSVFIVPIYIMACYSVVFCARROHALLNVNRHSLLEVVKD 240  
DB 181 LCSMIMGASPSYITLISVSVFIVPIYIMACYSVVFCARROHALLNVNRHSLLEVVKD 240  
QY 241 CVENEDEGAKEKKEEFODESEFRROHEGEVAKRGREAKDGLAKAEGSTGTSSEVYA 300  
DB 241 CVENEDEGAKEKKEEFODESEFRROHEGEVAKRGREAKDGLAKAEGSTGTSSEVYA 300  
QY 301 RGSEEVESSTVASDGSMEKKEGSKTEVENSMDKADKRTVEVNOCSIDLGEDDMFEGDDI 360  
DB 301 RGSEEVESSTVASDGSMEKKEGSKTEVENSMDKADKRTVEVNOCSIDLGEDDMFEGDDI 360  
QY 361 NFSEDDVEAVNIPESLPPSRNSNSNPDLPRCYOCKAKAVIIFIIISYVLSLGPYCLAV 420  
DB 361 NFSEDDVEAVNIPESLPPSRNSNSNPDLPRCYOCKAKAVIIFIIISYVLSLGPYCLAV 420  
QY 421 LAVWVDETVQPMVITIIIMLFLOCCIHPIYVGYGMHKTIKKEIDMLKKFCKEKRPK 480  
DB 421 LAVWVDETVQPMVITIIIMLFLOCCIHPIYVGYGMHKTIKKEIDMLKKFCKEKRPK 480  
QY 481 EDSDPDLPGTEGTEGKIVPSYDSATPP 508  
DB 481 EDSDPDLPGTEGTEGKIVPSYDSATPP 508

RESULT 4  
AAG64126  
ID AAG64126 standard; Protein; 508 AA.  
AC AAG64126;  
DT 25-SEP-2001 (first entry)  
DE Human G-protein-coupled receptor GPRV72.  
XX  
XX Human; guanosine triphosphate binding protein-coupled receptor;  
KW G-protein-coupled receptor; GPRV8; GPRV12; GPRV16; GPRV21; GPRV40;  
KW GPRV47; GPRV51; GPRV71; GPRV72; cancer; liver cirrhosis;  
KW Alzheimer's disease; cytosolic; hepatocellular; neurotrophic;  
KW neuroprotective; gene therapy; peptide therapy.  
XX  
OS Homo sapiens.  
XX



Query Match 99.9%; Score 2641; DB 22; Length 926;  
 Best local Similarity 99.8%; Pred. No. 1.5e-233;  
 Matches 507; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTCTNSTRESNSSHTCMPLSKMPISLAHGIIIRSTVLVIFLAASFVGNIVLALVLRKP 60  
 PS ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 1 MSTCTNSTRESNSSHTCMPLSKMPISLAHGIIIRSTVLVIFLAASFVGNIVLALVLRKP 60

QY 61 QLLQVTRNRFNFNLVLTDLQISLVAAPVAVTSVPLFPLNSHFTCTALVSLTHLFAFASVN 120  
 DB 61 QLLQVTRNRFNFNLVLTDLQISLVAAPVAVTSVPLFPLNSHFTCTALVSLTHLFAFASVN 120

QY 121 TIVLVSVDRYLSIIHPLSPYSKMTQRRGYLLTGTVAILQSTPPLYGMOAFAFDERNA 180  
 DB 121 TIVLVSVDRYLSIIHPLSPYSKMTQRRGYLLTGTVAILQSTPPLYGMOAFAFDERNA 180

QY 181 LCSMIMGASPSYTLISVSVFTVPLIYMIACYSVFCARROHALLYNKRHSLEVRKD 240  
 DB 181 LCSMIMGASPSYTLISVSVFTVPLIYMIACYSVFCARROHALLYNKRHSLEVRKD 240

QY 241 CVENEDDEGAKEKEEFODESEFRQHGHEVAKAKGRMEAKDGLAKEGSTGTSSESVEA 300  
 DB 241 CVENEDDEGAKEKEEFODESEFRQHGHEVAKAKGRMEAKDGLAKEGSTGTSSESVEA 300

QY 301 RGSEEVRESSSTVASDGSMEGKEGSTKYEENSMKADKGRTEVNOCSIDLGEDMFEGBDI 360  
 DB 301 RGSEEVRESSSTVASDGSMEGKEGSTKYEENSMKADKGRTEVNOCSIDLGEDMFEGBDI 360

QY 361 NFSEDDEAVANIPESLPPSRNSNSNPPLPRCYOCKAKAVIFIIIFSVYLSLGPYCFEAV 420  
 DB 361 NFSEDDEAVANIPESLPPSRNSNSNPPLPRCYOCKAKAVIFIIIFSVYLSLGPYCFEAV 420

QY 421 LAAWVDVETQVQWVITIIILFLLQCCIHPRYYGYMHTIKKEIDMLKKFCKEKP 480  
 DB 421 LAAWVDVETQVQWVITIIILFLLQCCIHPRYYGYMHTIKKEIDMLKKFCKEKP 480

QY 481 EDSHPDLPGTEGGTEGKIVPSYDSATPP 508  
 DB 481 EDSHPDLPGTEGGTEGKIVPSYDSATPP 508

RESULT 6  
 AAB61982  
 ID AAB61982 standard; protein; 508 AA.  
 AC AAB61982;  
 XX  
 DT 14-MAY-2001 (first entry)  
 XX  
 DE Human G-protein coupled receptor, IGSL.  
 XX  
 KW G-protein coupled receptor; IGSL; psychiatric; central nervous system;  
 movement disorder; tremor; Tourette's syndrome; Parkinson's disease;  
 Huntington's disease; dyskinesia; dystonia; spasm; neuroleptic; human;  
 neurotoxic; anticonvulsant; relaxant; vaccine; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200109184-A1.  
 XX  
 PD 08-FEB-2001.  
 XX  
 PF 17-JUL-2000; 2000WO-EP06678.  
 XX  
 PR 15-JUL-1999; 99EP-0202326.  
 PR 15-JUL-1999; 99NL-1012611.  
 XX  
 PA (SOLV ) SOLVAY PHARM BV.  
 XX  
 PI Deleersnijder W, Nys G, Zhang F;  
 XX WPI; 2001-182942/18.  
 DR N-PSDB; AAF56818.

XX  
 PT Novel human G-protein coupled receptor family polypeptide, IGSL, useful  
 PT for treating psychiatric and central nervous system disorders such as  
 PT tics, tremor, Tourette's syndrome and Parkinson's disease -  
 XX  
 PS Claim 15; Page 7; 65pp; English.  
 XX  
 CC This represents a G-protein coupled receptor family polypeptide, IGSL.  
 CC The IGSL protein can be expressed by standard recombinant methodology.  
 CC IGSL is useful for inducing immunological response in a mammal, as  
 CC vaccine. IGSL polynucleotides and polypeptides and its modulators are  
 CC useful for treating psychiatric and central nervous system disorders  
 CC especially for treating disorders, such as tics, tremor, Tourette's syndrome,  
 CC Parkinson's disease, Huntington's disease, dyskinesias, dystonia and  
 CC spasms.  
 XX  
 SQ Sequence 508 AA;

Query Match 99.7%; Score 2637; DB 22; Length 508;  
 Best local Similarity 99.8%; Pred. No. 1.5e-233;  
 Matches 507; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTCTNSTRESNSSHTCMPLSKMPISLAHGIIIRSTVLVIFLAASFVGNIVLALVLRKP 60  
 PS ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 1 MSTCTNSTRESNSSHTCMPLSKMPISLAHGIIIRSTVLVIFLAASFVGNIVLALVLRKP 60

QY 61 QLLQVTRNRFNFNLVLTDLQISLVAAPVAVTSVPLFPLNSHFTCTALVSLTHLFAFASVN 120  
 DB 61 QLLQVTRNRFNFNLVLTDLQISLVAAPVAVTSVPLFPLNSHFTCTALVSLTHLFAFASVN 120

QY 121 TIVLVSVDRYLSIIHPLSPYSKMTQRRGYLLTGTVAILQSTPPLYGMOAFAFDERNA 180  
 DB 121 TIVLVSVDRYLSIIHPLSPYSKMTQRRGYLLTGTVAILQSTPPLYGMOAFAFDERNA 180

QY 181 LCSMIMGASPSYTLISVSVFTVPLIYMIACYSVFCARROHALLYNKRHSLEVRKD 240  
 DB 181 LCSMIMGASPSYTLISVSVFTVPLIYMIACYSVFCARROHALLYNKRHSLEVRKD 240

QY 241 CVENEDDEGAKEKEEFODESEFRQHGHEVAKAKGRMEAKDGLAKEGSTGTSSESVEA 300  
 DB 241 CVENEDDEGAKEKEEFODESEFRQHGHEVAKAKGRMEAKDGLAKEGSTGTSSESVEA 300

QY 301 RGSEEVRESSSTVASDGSMEGKEGSTKYEENSMKADKGRTEVNOCSIDLGEDMFEGBDI 360  
 DB 301 RGSEEVRESSSTVASDGSMEGKEGSTKYEENSMKADKGRTEVNOCSIDLGEDMFEGBDI 360

QY 361 NFSEDDEAVANIPESLPPSRNSNSNPPLPRCYOCKAKAVIFIIIFSVYLSLGPYCFEAV 420  
 DB 361 NFSEDDEAVANIPESLPPSRNSNSNPPLPRCYOCKAKAVIFIIIFSVYLSLGPYCFEAV 420

QY 421 LAAWVDVETQVQWVITIIILFLLQCCIHPRYYGYMHTIKKEIDMLKKFCKEKP 480  
 DB 421 LAAWVDVETQVQWVITIIILFLLQCCIHPRYYGYMHTIKKEIDMLKKFCKEKP 480

QY 481 EDSHPDLPGTEGGTEGKIVPSYDSATPP 508  
 DB 481 EDSHPDLPGTEGGTEGKIVPSYDSATPP 508

RESULT 7  
 AAU04384  
 ID AAU04384 standard; protein; 508 AA.  
 AC AAU04384;  
 XX  
 DT 23-OCT-2001 (first entry)  
 XX  
 DE Human G-protein coupled receptor, hRUP15, mutant A398K.  
 XX  
 KW Human; G-protein coupled receptor; GPCR; hRUP15; agonist;  
 inverse agonist; lung cancer; A398K; mutant; mutein.  
 XX  
 OS Homo sapiens.

```

XX  MO200136471-A2.
PN
XX
PD  25-MAY-2001.
XX
PF  16-NOV-2000; 2000MO-US31509.
XX
PR  17-NOV-1999; 9905-0166088.
PR  17-NOV-1999; 9905-0166099.
PR  17-NOV-1999; 9905-0166369.
PR  23-DEC-1999; 9905-0171900.
PR  23-DEC-1999; 9905-0171901.
PR  23-DEC-1999; 9905-0171902.
PR  11-FEB-2000; 200005-0181749.
PR  14-MAR-2000; 200005-0189258.
PR  14-MAR-2000; 200005-0189259.
PR  10-APR-2000; 200005-0195898.
PR  10-APR-2000; 200005-0195899.
PR  10-APR-2000; 200005-0196078.
PR  28-APR-2000; 200005-0200419.
PR  12-MAY-2000; 200005-0203630.
PR  12-JUN-2000; 200005-0210741.
PR  12-JUN-2000; 200005-0210982.
PR  21-AUG-2000; 200005-0226760.
PR  26-SEP-2000; 200005-0235418.
PR  26-SEP-2000; 200005-0235779.
PR  20-OCT-2000; 200005-0242332.
PR  20-OCT-2000; 200005-0242343.
XX
XX  (AREN-) ARENA PHARM INC.
XX
XX  Chen R, Dang HT, Lowitz KP:
XX
XX  WPI: 2001-355616/37.
XX  N-PSDB; AAS08265.
XX
PT  Endogenous and non-endogenous versions of human G-protein coupled
PT  receptors for direct identification of candidate compounds as agonists,
PT  inverse agonists or partial agonists for use as therapeutic agents -
XX
XX  Claim 30; Page 141; 160pp; English.
XX
XX  The sequence is a human G-protein coupled receptor (GPCR), hRUP15, mutant
XX  A398K. The endogenous and non-endogenous, constitutively activated
XX  versions of human G-protein coupled receptors (GPCR), are useful for
XX  direct identification of candidate compounds as receptor agonists.
XX  Inverse agonists or partial agonists having applicability as therapeutic
XX  agents for treating diseases related to GPCR, e.g. lung cancer.
XX  Non-endogenous version of human GPCRs are also utilized in research
XX  settings and in vitro and in vivo system, incorporating GPCRs can be
XX  utilized to elucidate and understand the roles these receptors
XX  play in the human condition, both normal and diseased.
XX
XX  Sequence 508 AA:
XX
XX  Query Match 99.7%; Score 2636; DB 22; Length 508;
XX  Best Local Similarity 99.6%; Pred. No. 1,9e-233;
XX  Matches 506; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
QY  1 MTSTCTSTRESSNSHTCMPLSKMPLISLAGITRSTVLYFLAASFGNIVLALVLRKP 60
DB  1 MTSTCTSTRESSNSHTCMPLSKMPLISLAGITRSTVLYFLAASFGNIVLALVLRKP 60
QY  61 QLLQVTRRFTNLLVPTLLQISLVAAPVVAATSVPLFWPLNSHCTATVSLTHFAFASVN 120
DB  61 QLLQVTRRFTNLLVPTLLQISLVAAPVVAATSVPLFWPLNSHCTATVSLTHFAFASVN 120
QY  121 TIVLVSVDRYLSITHTPLSYPSKMTORRGYLLTGTVAILQSTPPYLGWGOAFAFDERNA 180
DB  121 TIVLVSVDRYLSITHTPLSYPSKMTORRGYLLTGTVAILQSTPPYLGWGOAFAFDERNA 180
QY  181 LCSMIMGASPSYTLISVVSFTIVPLIYMIACYSVVFCAARRQHALLVNKRHSLEVRKD 240
DB  181 LCSMIMGASPSYTLISVVSFTIVPLIYMIACYSVVFCAARRQHALLVNKRHSLEVRKD 240

```

```

DB  181 LCSMIMGASPSYTLISVVSFTIVPLIYMIACYSVVFCAARRQHALLVNKRHSLEVRKD 240
QY  241 CVENEDDEGAKEKKEFODESEFFRQHEGEYAKGEMKADGSLKAKEGSTGTSSESSVEA 300
DB  241 CVENEDDEGAKEKKEFODESEFFRQHEGEYAKGEMKADGSLKAKEGSTGTSSESSVEA 300
QY  301 RGSEVERESSSTVASDGSMEKKEGSTKVEENSMKADKRGTEVNOCSIDLGEDDMFEGDDI 360
DB  301 RGSEVERESSSTVASDGSMEKKEGSTKVEENSMKADKRGTEVNOCSIDLGEDDMFEGDDI 360
QY  361 NFSEDDYEAIVNIPESLPPSRNSNSNPPLPRCYCKAKAVFIITFSYVLSLGPYCLAV 420
DB  361 NFSEDDYEAIVNIPESLPPSRNSNSNPPLPRCYCKAKAVFIITFSYVLSLGPYCLAV 420
QY  421 LAVVVDVETQVPOKVTITITLFFLOCCIHPIYVGYVMHTKTKEIDOMLKKFCKEKP 480
DB  421 LAVVVDVETQVPOKVTITITLFFLOCCIHPIYVGYVMHTKTKEIDOMLKKFCKEKP 480
QY  481 EDSHPDLPGTEGTEGKIVPSYDSATPP 508
DB  481 EDSHPDLPGTEGTEGKIVPSYDSATPP 508

RESULT 8
AAU25561
ID  AAU25561 standard; Protein: 204 AA.
XX
XX  AAU25561:
XX
XX  18-DEC-2001 (first entry)
XX
XX  Human G Protein-Coupled Receptor (GPCR) polypeptide #8.
XX
XX  Human: G-protein coupled receptor; GPCR: mental disorder; schizophrenia;
XX  attention deficit disorder; anxiety; depression; bipolar disorder;
XX  neurological disorder; Huntington's disease; dementia; obesity; anorexia;
XX  type 2 diabetes; cardiovascular disorder; myocardial infarction; cancer;
XX  cardiomyopathy; atherosclerosis; human immunodeficiency virus; HIV;
XX  viral infection; immunostimulant; neuroleptic; nootropic; tranquiliser;
XX  antidepressant; anorectic; gene therapy.
XX
XX  Homo sapiens.
XX
XX  WO200162797-A2.
XX
XX  30-AUG-2001.
XX
XX  23-FEB-2001; 2001WO-US05676.
XX
XX  23-FEB-2000; 200005-0184247.
XX  23-FEB-2000; 200005-0184303.
XX  23-FEB-2000; 200005-0184304.
XX  23-FEB-2000; 200005-0184305.
XX  23-FEB-2000; 200005-0184397.
XX  02-MAR-2000; 200005-0186457.
XX  03-MAR-2000; 200005-0186810.
XX  09-MAR-2000; 200005-0188064.
XX  13-MAR-2000; 200005-0188880.
XX  03-APR-2000; 200005-0194344.
XX  23-JUN-2000; 200005-0213861.
XX  11-JUL-2000; 200005-0217369.
XX  11-JUL-2000; 200005-0217370.
XX  14-JUL-2000; 200005-0218337.
XX  20-JUL-2000; 200005-0218492.
XX
XX  (PHAA ) PHARMACIA & DUPHON CO.
XX
XX  Vogeli G, Wood LS, Parodi LA, Land P:
XX
XX  WPI: 2001-570628/64.
XX  N-PSDB; AAS42813.
XX

```



PT New isolated nucleic acid encoding a new G-protein coupled receptor  
PT polypeptide for detecting receptor modulators that can treat mental  
PT disorders, such as schizophrenia, anxiety, depression, or obesity -  
XX  
PS Claim 35; Page 73; 279pp; English.  
XX  
CC Sequences AAU2554-AAU25616 represent human G-protein coupled receptor  
CC (GPCR) polypeptides of the invention. The proteins and their associated  
CC DNA sequences can be used to identify compounds which bind to GPCR  
CC polypeptides and in screening for compounds that modulate GPCR activity.  
CC By screening a human subject for the presence of mutations in GPCR DNA, a  
CC GPCR-related disorder or a genetic predisposition can be diagnosed. The  
CC sequences can also be used for treatment and prevention of mental  
CC disorders such as schizophrenia, attention deficit disorder, anxiety,  
CC depression, dementia and bipolar disorder, neurological disorders such as  
CC Huntington's disease, Parkinson's disease and Tourette's syndrome,  
CC metabolic disorders such as obesity, anorexia and type 2 diabetes,  
CC cardiovascular disorders such as thrombosis, myocardial infarction,  
CC cardiomyopathy and atherosclerosis, viral infections caused by HIV and  
CC cancers.  
CC  
XX  
SQ Sequence 204 AA:  
Query Match 29.9%; Score 791; DB 22; Length 204;  
Best Local Similarity 98.7%; Pred. No. 1.6e-64;  
Matches 156; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MISTCTNSTRESNSHTCMPLSKMPISLANGITRSTVLTFLAASFGNVLALVLRKP 60  
DB 47 MISTCTNSTRESNSHTCMPLSKMPISLANGITRSTVLTFLAASFGNVLALVLRKP 106  
QY 61 QLLQVTRFNFNLVTDLLQISLVAPWVATSVPLFWPLNSHCTALVSLTHLFAFASV 120  
DB 107 QLLQVTRFNFNLVTDLLQISLVAPWVATSVPLFWPLNSHCTALVSLTHLFAFASV 166  
QY 121 TIVLVSVDRYLSIIHPLSYPSKMTORRGYLLLTWIV 158  
DB 167 TIVLVSVDRYLSIIHPLSYPSKMTORRGYLLLTWIV 204  
RESULT 9  
AA70343  
ID AA70343 standard; Protein; 407 AA.  
XX  
AC AA70343;  
XX  
DT 20-JUN-2000 (first entry)  
XX  
DE Human G protein-coupled orphan receptor, RE2.  
XX  
KM G protein-coupled orphan receptor; GPCR; agonist; G protein; treatment;  
KM GPCR fusion protein; Inverse agonist; drug; RE2; human.  
XX  
OS Homo sapiens.  
XX  
PN WO200006597-A2.  
XX  
PD 10-FEB-2000.  
XX  
PF 30-JUL-1999; 99WO-US17425.  
XX  
PR 31-JUL-1998; 98US-0094879.  
PR 30-OCT-1998; 98US-0106300.  
PR 04-DEC-1998; 98US-0110906.  
PR 26-FEB-1999; 99US-0121851.  
XX  
PA (AREN-) ARENA PHARM INC.  
XX  
PI Behan DP, Chalmers DT, Liaw C, Lin I, Lowitz K, Chen R;  
XX  
DR WPI; 2000-195260/17.  
XX  
DR N-PSDB; AA551461.  
XX

PT Identification of a compound useful as a therapeutic agent, comprises  
PT identifying a compound against constitutively activated G  
PT protein-coupled orphan receptors -  
XX  
PS Example 2; Page 111; 123pp; English.  
XX  
CC The patent discloses a method of identifying agonists and inverse or  
CC partial agonists to the endogenous, constitutively activated  
CC G protein-coupled orphan receptors (GPCRs), by contacting them with a  
CC GPCR fusion protein comprising a GPCR and a G protein. Determining  
CC expression of GPCRs in tissue samples can be used to identify related  
CC diseases. Inverse agonists to these receptors can be used as drugs for  
CC treating GPCR-related diseases. The present protein sequence is that of  
CC human G protein-coupled orphan receptor, RE2.  
XX  
SQ Sequence 407 AA:  
Query Match 14.8%; Score 391; DB 21; Length 407;  
Best Local Similarity 23.6%; Pred. No. 2.3e-27;  
Matches 106; Conservative 67; Mismatches 130; Indels 146; Gaps 11;  
QY 31 GIIRSTVLTFLAASV--GNIVLALVLRKPQLQVTRFNFNLVTDLLQISLVAPW 88  
DB 25 GVITTFIATIVITIFVCLGNLIVVTLTKKSYLLTSLNKFVSLTSLNLFSLVLPFV 84  
QY 89 VMTSVPLFWPLNSHCTALVSLTHLFAFASVNTIVLVSVDRYLSIIHPLSYPSKMTORRG 148  
DB 85 VTSSIRREWFNGVWCMFSAALYLLISSASMLTGVAIDRYAVLVPMVPMKKTGNRA 144  
QY 149 YLLVYGTWIAVILQSTPLPLGWSGOAFDERNALCSMTGASPSYTLISVSVFIPLIV 208  
DB 145 VMAVLYITMLSHLQCLPLPLGWSVSEDFERKMCVAMHREPGTAWQIMCALFPLIV 204  
QY 209 IACYSVFCAARHALLVYVKKRSHLEVRVYKDCVENEDEGAEKKEEFODESEFRROHEG 268  
DB 205 IACYGFEYFVAR-----VKARKVHCGYVIVE-BDAQ----- 235  
QY 269 EVKAKKEGRMEKQSLAKKSGTSTSSVSEARSESEVRSSYVASGSMCKGKSTKYE 328  
DB 236 ---RTGR-----KNSSTSTSSG----- 250  
QY 329 ENSMKADKGRTEVNOCSIDLGEDDMERGEDDINFSEDDVAVNIPESLPSRRNS----- 383  
DB 251 -----SRRNAPGVV 260  
QY 384 -NSNPPLRCYQCKAAVFIILIFSIVLSLGPY-CFLAVLAW--VDVETQVPMVITTI 439  
DB 261 YSAN-----QCKALITILVLGAFVWTGMPYVWVIASEALMCKSSVSPSLFTWA----- 309  
QY 440 TWLFLOCCHIPPYYGYGMHTIKKEIODM 468  
DB 310 TWLSEFASAVCHPLIYGLMNTVRKELLGM 338  
RESULT 10  
AAW58586  
ID AAW58586 standard; Protein; 454 AA.  
XX  
AC AAW58586;  
XX  
DT 07-SEP-1998 (first entry)  
XX  
DE Human histamine H2 receptor.  
XX  
KM Human; histamine H2 receptor; H2RH; inflammatory disease; gastric;  
KM nervous condition.  
XX  
OS Homo sapiens.  
XX  
PI WO9820040-A1.  
XX  
PN WO9820040-A1.  
XX  
PD 14-MAY-1998.  
XX

```

FF      05-NOV-1997;       97WO-US20200.
XX
XX      07-NOV-1996;       96US-0748485.
XX
PA      (INCY-) INCYTE PHARM INC.
XX
XX      Au-Young J, Goli SK, Guegler KJ, Murry LE;
PI      WPI: 1998-286870/25.
DR      N-PDSB: AAV37701.
PT
PT      Histamine H2 receptor - used to treat inflammatory disease, gastric
XX      conditions and nervous conditions
XX
XX      Claim 1, Page 51-52; 77pp; English.
CC
CC      The present sequence represents histamine H2 receptor (H2RH) isolated
CC      from Incyte Clone 1722180 from the human bladder cDNA library
CC      (BLADNDT05). A host cell, comprising a vector which contains the nucleic
CC      acid encoding H2RH, can be used to produce the H2RH. An antibody which
CC      specifically binds the H2RH can be used to detect and quantify H2RH in
CC      a biological sample. An antagonist which specifically binds to and
CC      modulates the activity of H2RH can be used in a pharmaceutical
CC      composition for treating inflammatory disease, gastric conditions, and
CC      nervous conditions. The gastric conditions that can be treated using the
CC      antagonist, include gastritis, flu, colitis, and Crohn's disease. The
CC      nervous conditions include Alzheimer's disease, ataxia, Eaton-Lambert
CC      syndrome, epilepsy, myasthenia gravis, and Parkinson's disease. The
CC      antagonist may also be used to treat infections or inflammation of the
CC      urinary tract and bladder. It may also be used to modulate H2RH activity
CC      in endothelial cells of the cardiovascular system and treat diseases
CC      such as arteriosclerosis, cardiomyopathy, endocarditis, and ischemia.
XX
XX      Sequence    454 AA;
SQ
          Query Match           14.7%; Score 388.5; DB: 19; Length 454;
          Best Local Similarity   22.9%; Pred. No. 4.5e-27;
          Matches 103; Conservative 72; Mismatches 123; Indels 145; Gaps 10.
OY      31 GIIRSTVFVIFLAASFV--GNIVLALVLQRRKOLLQVTNRFEFNLVYDLDLOISLVAPMV 88
Db      25 GVIIHQFIAYITTFEVCGLNVIYVTVTKKSYLLTSNKPFVSLTSLNFLLSVLYLPEV 84
OY      89 VATSPLEWPLNSHCCTALVSLTHLFARASVTVILVYSVDRLSTIHPYSTSKMTQRG 148
Db      85 VTSSIRREMIEGVWCNFSALLYLLISSASMLTGIALIDRYAVALPYMVKTIIGNRA 144
OY      149 YLLFGVIVIALOSTPPLYWGQAFAFDPRNALGSMIGCASPVYTLVSFTVPILYM 208
Db      145 VMALVYIMWLHSILGCLPLFGMSYVEPDFEKMKCAVAHIREGYTAFMQICALPPFLYM 204
OY      209 IACYSVVFCARQHALLYNKRHSLEVVRKDCVENEDEGAKEKEEFODESEFRRQHEG 268
Db      205 LWCVGYEFFRVAR----- 216
OY      269 EFKAKEGRNEMADGSLKKKEGSTGTSESSVEARGSEVEESTVASDGMECKEGSTKYE 328
Db      217 -VKAR-----KHHGGTVYLVEDDAQRTG--VKNKSSTSTSSS----- 250
OY      329 ENSMAKDGRTEVNOCSIDLGEDDEFEDEDINFSEDDVEAVNIPELSPPSRNRS---- 383
Db      251 -----GSRRNAFGQVV 261
OY      384 -NSNPPLPRCYCCKAAKYVITITTFESVLSLGRY-CFLAVLANV--VDVETQVPQWYITII 439
Db      262 YSAN-----OCKALITLVLAGFMVGMVGMYAVIASALMGKSSVPSLETWA--- 310
OY      440 TWFELQCCIHPYVYGVMKTIKKEIODM 468
Db      311 TWLSFAVAACHPLITIGLMNKTYRKELLGM 329

```

[illegible]

```

Db 274 ----- 273
Oy 364 EDDVEAVNIPESLPSPRRNS-----NSNPPLPCYQCKAAKVFIIIFSYLSIGPY-C 416
Db 274 -----SRNAFGGVVYSAN-----QCKALITLIVLGAFVWTGWPYV 311
Oy 417 FLAVLAWV--VDVEQVQVWVITIIIMLFLLQCCIHPIVYGYMHTIKKEIQDM 468
Db 312 VIASEALMGKSSVSPSLETMA---TWLSFASAVCHSLIYGLMNTVKRELLGM 361

RESULT 12
AAR90989
ID AAR90989 standard; Protein: 529 AA.
XX
AC AAR90989;
XX
XX 27-MAY-1996 (first entry)
XX
DE Human adrenergic G-protein coupled receptor.
XX
XX G-protein coupled receptor polypeptide; GCRP; adrenergic receptor;
XX agonist; antagonist; therapy; hypertension; respiratory disease.
XX
XX Homo sapiens.
XX
XX MO9605225-A1.
XX
XX 22-FEB-1996.
XX
XX 10-AUG-1994; 94WO-US09051.
XX
XX 10-AUG-1994; 94WO-US09051.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX PA
XX PI Adams MD, LI Y, Soppet DR;
XX
XX WPI; 1996-139642/14.
XX
XX N-PSDB; AAT13002.
XX
XX New isolated G-protein coupled receptor polypeptide - used to
XX develop prods. to inhibit or stimulate adrenergic receptors for
XX treating e.g. hypertension or respiratory disorders
XX
XX Claim 4; Fig 1; 60pp; English.
XX
XX Human G-protein coupled receptor polypeptide (GCRP) (AAR90989)
XX is a protein structurally related to the alpha-1 adrenergic
XX receptor family. It can be produced by expression of a cDNA clone
XX (AAT13002) isolated from a human infant brain cDNA library.
XX Recombinant GCRP is used to screen for agonist and antagonist
XX cpts. useful e.g. in the treatment of hypertension and respiratory
XX disorders, and to raise anti-GCRP antibodies.
XX
XX Sequence 529 AA;

Query Match 14.0%; Score 371; DB 17; Length 529;
Best Local Similarity 23.1%; Pred. No. 2,2e-25;
Matches 110; Conservative 69; Mismatches 132; Indels 166; Gaps 14;

```

```

Oy 181 LCSMTMGASPSYTLISVYSFIYIPLIIMACYSVYECARBOHALLYNKRHSLEVRKD 240
Db 177 MCVAAMHREPGTTAEMQIMCALFPPLVIMVCGTIFRYAR-----VKARKHCGTIV 228
Oy 241 CVENDEDEGAKEKEEFODESEFRROHEGEVAKESGRMEAKDGLAKKEGSTGTSSSVEA 300
Db 229 IVE-EDAQ-----RIGR-----KNSSTSTSS----- 249
Oy 301 RGESEVRRESSIVASDGSMEGKEGCTVBEENSKADKRTENVQCSIDLGEDDMERGEDDI 360
Db 250 -----GR----- 251
Oy 361 NFSEDDVEAVNIPESLPSPRRNS-----NSNPPLPCYQCKAAKVFIIIFSYLSIGP 414
Db 252 -----RRNAFGGVVYSAN-----QCKALITLIVLGAFVWTGWP 285
Oy 415 Y-CFLAVLAWV--VDVEQVQVWVITIIIMLFLLQCCIHPIVYGYMHTIKKEIQDM 468
Db 286 YMWVVIASEALMGKSSVSPSLETMA---TWLSFASAVCHPLIYGLMNTVKRELLGM 338

RESULT 13
AAM70501
ID AAM70501 standard; Protein: 529 AA.
XX
XX AAM70501;
XX
XX 29-DEC-1998 (first entry)
XX
XX Human G-protein adrenergic receptor protein sequence.
XX
XX Human G-protein adrenergic receptor; baculovirus expression system;
XX COS cell; upper respiratory condition; hypertension.
XX
XX Homo sapiens.
XX
XX OS
XX PA
XX PI US5817477-A.
XX
XX PD 06-OCT-1998.
XX
XX PF 06-JUN-1995; 95US-0467568.
XX
XX PR 06-JUN-1995; 95US-0467568.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX PA
XX PI Adams MD, LI Y, Soppet DR;
XX
XX DR WPI; 1998-556386/47.
XX
XX N-PSDB; AAV33510.
XX
XX DNA encoding adrenergic receptor polypeptide - useful for producing
XX recombinant polypeptide and identifying antagonists and agonists of
XX the receptor
XX
XX Claim 1; Fig 1A-1E; 38pp; English.
XX
XX The present DNA sequence represents a human G-protein adrenergic receptor
XX protein sequence. The receptor is encoded by DNA isolated from a human
XX infant cDNA library. Vectors and host cells (e.g. Baculovirus expression
XX system, COS cells, etc.) can be used for recombinant production of the
XX receptor. The recombinant receptor is claimed to be useful for
XX identifying antagonists and agonists of the receptor, and for raising
XX antibodies against it. The agonists are claimed useful for treating
XX disease conditions associated with the receptor activity, e.g upper
XX respiratory conditions and hypertension.
XX
XX Sequence 529 AA;

Query Match 14.0%; Score 371; DB 19; Length 529;
Best Local Similarity 23.1%; Pred. No. 2,2e-25;
Matches 110; Conservative 69; Mismatches 132; Indels 166; Gaps 14;

```





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 16:45:08 ; Search time 16 Seconds

(without alignments)  
775.513 Million cell updates/sec

Title: US-09-992-238-2

Perfect score: 2644

Sequence: 1 MTSTCTNSTENSSHTCMP.....GTEGTEGKIVPSYDSATFP 508

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/1aa/5b.COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/1aa/PCtUS.COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	388.5	14.7	454	2	US-08-748-485-1
2	371	14.0	529	2	US-08-467-568-2
3	371	14.0	529	5	PCT-US94-09051-2
4	371	14.0	529	5	PCT-US94-09051-2
5	343.5	13.0	422	1	US-07-817-920-3
6	343.5	13.0	422	1	US-08-370-542-3
7	343.5	13.0	422	1	US-08-117-006-3
8	343.5	13.0	422	1	US-08-216-594-3
9	343.5	13.0	422	1	US-08-542-358-3
10	343.5	13.0	422	2	US-08-157-185-13
11	343.5	13.0	422	3	US-08-281-526B-13
12	343.5	13.0	422	3	US-09-018-351-3
13	343.5	13.0	422	5	PCT-US93-00149-3
14	338.5	12.8	466	1	US-08-334-698-6
15	338.5	12.8	466	1	US-08-228-932-6
16	338.5	12.8	466	1	US-08-468-939-6
17	338.5	12.8	466	1	US-08-722-001-28
18	338.5	12.8	466	2	US-08-406-855A-6
19	338.5	12.8	466	2	US-08-722-190-6
20	338.5	12.8	466	3	US-08-244-354-6
21	338.5	12.8	466	5	PCT-US95-04203-6
22	338.5	12.8	466	5	PCT-US95-04203-6
23	333.5	12.6	429	2	US-08-748-485-7
24	333.5	12.6	466	1	US-08-722-001-12
25	333.5	12.6	466	2	US-08-467-568-11
26	333.5	12.6	466	2	US-09-030-582-11
27	324.5	12.3	466	2	US-08-406-855A-23

28	324.5	12.3	466	3	US-09-206-899-23	Sequence 23, Appl
29	314.5	11.9	501	1	US-08-722-001-14	Sequence 14, Appl
30	314.5	11.9	501	2	US-08-467-568-9	Sequence 9, Appl
31	314.5	11.9	501	2	US-09-030-582-9	Sequence 9, Appl
32	314.5	11.9	572	1	US-08-334-698-2	Sequence 2, Appl
33	314.5	11.9	572	1	US-08-228-932-2	Sequence 2, Appl
34	314.5	11.9	572	1	US-08-468-939-2	Sequence 2, Appl
35	314.5	11.9	572	1	US-08-722-001-30	Sequence 30, Appl
36	314.5	11.9	572	2	US-08-406-855A-2	Sequence 2, Appl
37	314.5	11.9	572	2	US-08-722-190-2	Sequence 2, Appl
38	314.5	11.9	572	3	US-08-244-354-2	Sequence 2, Appl
39	314.5	11.9	572	3	US-09-206-899-2	Sequence 2, Appl
40	314.5	11.9	572	5	PCT-US95-04203-2	Sequence 2, Appl
41	309.5	11.7	521	2	US-08-406-855A-19	Sequence 19, Appl
42	309.5	11.7	521	3	US-09-206-899-19	Sequence 19, Appl
43	304.5	11.5	601	1	US-07-676-174A-2	Sequence 2, Appl
44	302.5	11.4	375	1	US-08-118-270-17	Sequence 17, Appl
45	302.5	11.4	375	5	PCT-US93-08528-17	Sequence 17, Appl

## ALIGNMENTS

RESULT 1  
US-08-748-485-1  
Sequence 1, Application US/08748485  
Patent No. 5817480  
GENERAL INFORMATION:  
APPLICANT: Au-Young, Janice  
APPLICANT: Guegler, Karl J.  
APPLICANT: Goli, Surya K.  
APPLICANT: Murry, Lynn E.  
TITLE OF INVENTION: NOVEL HISTAMINE H2 RECEPTOR  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/748,485  
FILING DATE: Herewith  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0159 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 454 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: Consensus  
CLONE: 1722180  
US-08-748-485-1  
Query Match 14.7%, Score 388.5, DB 2, Length 454;





TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ. ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 529 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-030-582-2

Query Match 14.0%; Score 371; DB 2; Length 529;  
Best Local Similarity 23.1%; Pred. No. 3.3e-22;  
Matches 110; Conservative 69; Mismatches 132; Indels 166; Gaps 14;

QY 6 TNSTRE-----SNSHTCMPLSKMPISLHGIIRSTVLVIFLASFGNYLALVLRKP 60  
DB 14 SNLTREEGEGASSPSSPSLSP-----FLSA--WGNLVIYVTLTKKS 56  
QY 61 QLLQVTRFTFNLVTDLDLSIAPVAVATSVPLFWPLNSHFTALVSLTLFAFASVN 120  
DB 57 YLLTSLKKEFVSLTSLNLSVLVLPVTVSSIRREMIPEGVWCMFSAALYLLISSASML 116  
QY 121 TIVLVSDRYLSIHPVPSYKMTORRGYLLLTGTWVAIIQSTPPLYGMOAFADERNA 180  
DB 117 TLGVIAIDRYAVLYPMVYPMKITGNRAVALYIWLHSLGCLPLFGSSVEYGENKW 176  
QY 181 LGSIMWASPSYTLISVSVFTIPLIYMIACYSVFCARQHALLYNKRHSLEVRKD 240  
DB 177 MCVAAHMRREGYTAFWQIMCALPEFLVLCYGFIFRVAR-----VARKVHCCTVV 228  
QY 241 CVENEDEGAKEKEEFODESEFRQHEGEVAKKGRMEAKDGLAKEGSTGTSSEVEA 300  
DB 229 IVE-EDAQ-----RTGR-----KNSSTSTSSS----- 249  
QY 301 RGSEEVRESSTVASDGSMEKSGSTKVEENSMKADGRTEVNCSDIDGEDMEFGEEDI 360  
DB 250 -----GR----- 251  
QY 361 NFEEDVEAVNIPESLPSPRRNS-----NSNPPLPRCYOCKAKAVIFIIIFSVLSLP 414  
DB 252 -----RRNAFGGVVYSAN-----OCKALITLIVLGAFAWVTWGP 285  
QY 415 Y-CFLAVLAWV--VDVETQVPQWVITIIIMLFLOCCIHPPVYGYMKHTIKKEIQDM 468  
DB 286 YMVVIASEALMGKSSVSPSLETWA---TWLSFASAVCHPLITGLMKNKYRKELLGM 338

RESULT 4  
PCT-US94-09051-2  
Sequence 2, Application PC/TUS9409051  
GENERAL INFORMATION:  
APPLICANT: LI, ET AL.  
TITLE OF INVENTION: Adrenergic Receptor  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/09051  
FILING DATE: Submitted herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: FERRARO, GREGORY D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-194  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ. ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 529 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS:  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
PCT-US94-09051-2

Query Match 14.0%; Score 371; DB 5; Length 529;  
Best Local Similarity 23.1%; Pred. No. 3.3e-22;  
Matches 110; Conservative 69; Mismatches 132; Indels 166; Gaps 14;

QY 6 TNSTRE-----SNSHTCMPLSKMPISLHGIIRSTVLVIFLASFGNYLALVLRKP 60  
DB 14 SNLTREEGEGASSPSSPSLSP-----FLSA--WGNLVIYVTLTKKS 56  
QY 61 QLLQVTRFTFNLVTDLDLSIAPVAVATSVPLFWPLNSHFTALVSLTLFAFASVN 120  
DB 57 YLLTSLKKEFVSLTSLNLSVLVLPVTVSSIRREMIPEGVWCMFSAALYLLISSASML 116  
QY 121 TIVLVSDRYLSIHPVPSYKMTORRGYLLLTGTWVAIIQSTPPLYGMOAFADERNA 180  
DB 117 TLGVIAIDRYAVLYPMVYPMKITGNRAVALYIWLHSLGCLPLFGSSVEYGENKW 176  
QY 181 LGSIMWASPSYTLISVSVFTIPLIYMIACYSVFCARQHALLYNKRHSLEVRKD 240  
DB 177 MCVAAHMRREGYTAFWQIMCALPEFLVLCYGFIFRVAR-----VARKVHCCTVV 228  
QY 241 CVENEDEGAKEKEEFODESEFRQHEGEVAKKGRMEAKDGLAKEGSTGTSSEVEA 300  
DB 229 IVE-EDAQ-----RTGR-----KNSSTSTSSS----- 249  
QY 301 RGSEEVRESSTVASDGSMEKSGSTKVEENSMKADGRTEVNCSDIDGEDMEFGEEDI 360  
DB 250 -----GR----- 251  
QY 361 NFEEDVEAVNIPESLPSPRRNS-----NSNPPLPRCYOCKAKAVIFIIIFSVLSLP 414  
DB 252 -----RRNAFGGVVYSAN-----OCKALITLIVLGAFAWVTWGP 285  
QY 415 Y-CFLAVLAWV--VDVETQVPQWVITIIIMLFLOCCIHPPVYGYMKHTIKKEIQDM 468  
DB 286 YMVVIASEALMGKSSVSPSLETWA---TWLSFASAVCHPLITGLMKNKYRKELLGM 338

RESULT 5  
US-07-817-920-3  
Sequence 3, Application US/07817920  
Patent No. 5360735  
GENERAL INFORMATION:  
APPLICANT: Welshank, Richard L  
APPLICANT: Branchek, Theresa  
APPLICANT: Hartig, Paul R  
TITLE OF INVENTION: DNA ENCODING A HUMAN 5-HT1F RECEPTOR AND  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/817,920
FILING DATE: 19920108
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 1795/39318
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-977-9550
TELEFAX: 212-664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 422 amino acids
TYPE: AMINO ACID
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
IMMEDIATE SOURCE:
CLONE: 5-HT1A
US-07-817-920-3

```

```

Query Match      13.0%; Score 343.5; DB 1; Length 422;
Best Local Similarity 22.0%; Pred. No. 4.1e-20;
Matches 101; Conservative 92; Mismatches 190; Indels 77; Gaps 9;

```

```

QY 26 ISLAHGIRSTVLVIFLAASFGN--IVLALVLRKPOLQVYTRFNFLLYTDLQISL 83
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 31 VTVSYOYITSLGLTFLFCVAGNACVVAIALER--SLQNVANYILGSLAVTDLMVSVL 88
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 84 VAPWVAVTSVPLFMPPLNSHCTALVSLTHLFAFASVNTIVLSVDRLSTIIHPLSPSKM 143
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 89 VLPMAALYQVLNKTGLQVYCDLFIALDVLCCTSSIIHLCAIALDRYWAITDPIDYVNR 148
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 144 TORRGYLLYGTWIVAILQSTPPLYGNGQAAPDERNALCMINGASPSYTIILSVSEFVI 203
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 149 TPRRAALLISLTWILGLFLLISIPPLGWRTP--EDRSDPDACCTISKDGYIYSTFGAFYI 206
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 204 PLIVIMACYSVVCAARQALALYNVRHSLLEVAVKDCVENEDEGAKEKEEPQDESEFR 263
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 207 PLLIMLVLYGRIFRAAR-----PRIKTYKVKVETGADTRHGASPAQPK 251
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 264 ROHEGEVAKAEGRM--EAKDGL-----KAKEGSTGTSSESSVEARGSEEVRESSTVASDG 316
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 252 KSYNGEGSGNMRIGVSKAGALCANGAVRQGDGALALEVIEVH----- 296
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 317 SMEGKGSTVVEENSKADKRGTEVNOCSIDLGEDMEFGEEDINSEDDVEAVNIPESL 376
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 297 ----RVGNRSK-EHLPLPSEAGPTPCAPASFE----- 322
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 377 PPSRRNSNSNPRLPCRCQCAAKVIFIIIFSYVLSLGPYCFGLAVLAVWVDEVQVPOWVI 436
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 323 RKNERNEARAKMALARERTVKTGLITMGTFILCWLPEFIVALLVLPFCSSSCHMPTLLG 382
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 437 TTIIMLFLOCCIHPIYVGYMHKTIKEIDMLKKFPCKE 476
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 383 AIIINMIGYSNLSLNPVYALFNFMDQNAFKIITKICLCFCRQ 422
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 6
US-08-370-542-3
; Sequence 3, Application US/08370542
; Patent No. 5476782
; GENERAL INFORMATION:
; APPLICANT: Weinsinank, Richard L.

```

```

APPLICANT: Branchek, Theresa
APPLICANT: Hartly, Paul R.
TITLE OF INVENTION: DNA ENCODING A HUMAN 5-HT1E RECEPTOR AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/370,542
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/194,113
FILING DATE:
APPLICATION NUMBER: US/07/803,626
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 1795/39317
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-977-9550
TELEFAX: 212-664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 422 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-370-542-3

```

```

Query Match      13.0%; Score 343.5; DB 1; Length 422;
Best Local Similarity 22.0%; Pred. No. 4.1e-20;
Matches 101; Conservative 92; Mismatches 190; Indels 77; Gaps 9;

```

```

QY 26 ISLAHGIRSTVLVIFLAASFGN--IVLALVLRKPOLQVYTRFNFLLYTDLQISL 83
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 31 VTVSYOYITSLGLTFLFCVAGNACVVAIALER--SLQNVANYILGSLAVTDLMVSVL 88
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 84 VAPWVAVTSVPLFMPPLNSHCTALVSLTHLFAFASVNTIVLSVDRLSTIIHPLSPSKM 143
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 89 VLPMAALYQVLNKTGLQVYCDLFIALDVLCCTSSIIHLCAIALDRYWAITDPIDYVNR 148
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 144 TORRGYLLYGTWIVAILQSTPPLYGNGQAAPDERNALCMINGASPSYTIILSVSEFVI 203
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 149 TPRRAALLISLTWILGLFLLISIPPLGWRTP--EDRSDPDACCTISKDGYIYSTFGAFYI 206
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 204 PLIVIMACYSVVCAARQALALYNVRHSLLEVAVKDCVENEDEGAKEKEEPQDESEFR 263
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 207 PLLIMLVLYGRIFRAAR-----PRIKTYKVKVETGADTRHGASPAQPK 251
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 264 ROHEGEVAKAEGRM--EAKDGL-----KAKEGSTGTSSESSVEARGSEEVRESSTVASDG 316
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 252 KSYNGEGSGNMRIGVSKAGALCANGAVRQGDGALALEVIEVH----- 296
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 317 SMEGKGSTVVEENSKADKRGTEVNOCSIDLGEDMEFGEEDINSEDDVEAVNIPESL 376
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 297 ----RVGNRSK-EHLPLPSEAGPTPCAPASFE----- 322
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db      149  TPRRAAAILSLTWLIGFLISIPRLGWRTP--EDRSDPADACTISKDHGYITTSFGAFYI 206
Qy      204  PLIYMIACISVYFCAARQNALNVKRNHSLFNVAKCVCENEDEGAEKKEEFODESEFR 253
Db      207  PLLLMLVLYGRIFFRAA-----FIRRTVKKVEGTGADTFRHGASPAQPK 251
Qy      264  RQHGEVAKKGRN--FAKDGL-----KKEGSTGSESSVEARGSEVRESSTVASDG 316
Db      252  KSVNGESSRWRLGVESKAGALCANGAVRQGDGALEVEIYH----- 296
Qy      317  SMEKEGSTKYEENSMAKDGRTGVNOCSIDLGEDMEFGEDDINFSEDDVEAVNIPESL 376
Db      297  ----RVGNSK-EHLPLDSEAGPTFCAPASF----- 322
Qy      377  PPSRRNSNPNPLDPRCYOAKAYIFIIISVLSLGPYCFGLAVLAVWDVETQVPQWT 436
Db      323  RKNREINAEAKRKMLALAREKRTVKTLGIMGFILCWLPEFVALVLPFCSSCHMPTLLG 382
Qy      437  TIIIMLPLOCCIHPIYVGYMHNKTIKKEIDMLKKFFCKE 476
Db      383  AINMLGYSNLSLNPVIYAVFNKDFQNAKFIKICLCRCQ 422

RESULT 8
US-08-216-594-3
; Sequence 3, Application US/08216594
; Patent No. 5652113
; GENERAL INFORMATION:
; APPLICANT: Weinschank, Richard L.
; APPLICANT: Branchek, Theresa
; APPLICANT: Hartlig, Paul R.
; TITLE OF INVENTION: DNA ENCODING A HUMAN 5-HT1F RECEPTOR
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/216,594
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 1795/39318
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-977-9550
; TELEFAX: 212-664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 422 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; IMMEDIATE SOURCE:
; CLONE: 5-HT1A
US-08-216-594-3

```



```

;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible

```

; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Cooper & Dunham LLP  
 ; STREET: 1185 Avenue of the Americas  
 ;

```

: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.24
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/281,526B
: FILING DATE: 27-JUL-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: White P., John
: REGISTRATION NUMBER: 28,678
: REFERENCE/DOCKET NUMBER: 41908-1/JPW
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 278-0400
: TELEFAX: (212) 391-0525
: INFORMATION FOR SEQ ID NO: 13:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 422 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-281-526B-13
:
: Query Match 13.0%; Score 343.5; DB 3; Length 422:
: Best Local Similarity 22.0%; Pred. No. 4,1e-20:
: Matches 101; Conservative 92; Mismatches 190; Indels 77; Gaps
:
: Oy 26 ISLAGIIRSTVLVIFLASFVGN--IVLALVLRKKPOLLOYTNRPIFLMLVTDLQISL 83
:   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
: Db 31 VIVSVQVITSLILGTLIFCAVLGNACVVAIALDER--SLQNVANYLIGSLAVTDLWVSL 88
:   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
: Oy 84 VAPWVAIVSVPLFVPLNSHFTALVSLTHLFAFASVNTIVLVSVDRIYLSIHPLSPSKM 143
:   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
: Db 89 VLPMAALYQVLNKKWTLGQVTCDLFIADLVLCCTSSILHLCAIALDRYMAITDPIDVKNR 148
:   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
: Oy 144 TORRGILLGYWVAIILQSTPEPLVGMGOAEDERNALCSMTMGASPTILSVSFIVY 203
:   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
: Db 149 TPRRAALISLTLWILGFLISIPMLGWRR--EDNSDPDCAITISKDHGTTIYSTFGAFYI 206
:   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
: Oy 204 PLIVNIACYSVFECARROHALLVYVNRHSLEVRVDCVENEDEGAERKEEFODESEFR 263
:   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
: Db 207 PLLMLVLVYGRIFRAAR-----FRIKRTVKVKEKTGADTRHGASPAPOPK 251
:   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
: Oy 264 ROHESEVAKKEGRM--EAKDGL-----KAKGGSTGTSSESVYANGSEVEYRESSIVASDG 316
:   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
: Db 252 KSVNEGSEGRNMRLVYESKAGALCANGAVRQGDGAALEVIEVH----- 296
:   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
: Oy 317 SMEGREGSTKVEENSKAKAKRTVEVNOCSIDLGEDDMERGEDDINSEDDVEAVNIPESL 376
:   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
: Db 297 ---RVGNSK-DHLPLPSAGPTPCAPASFE----- 322
:   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
: Oy 377 PSRRNSNSNPDLPRCYOCRAKAKVFIILFFSVLSLGPFCFLAVLAWVDVETPOWVI 436
:   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
: Db 323 RKNENNAEKRRKMLAREKTYKTIGTIINGTITLCLMPFIYALVLPCESSCHMPTLLG 382
:   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
: Oy 437 TIIWFLQCCIHPEVYGYVMKTIKKEIQMLKKEFFE 476
:   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
: Db 383 AIIWNLGYSNLSLNPVIYAYFNKKDFQONAPKRIIKCLFCHQ 422
:   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
:
: RESULT 12
: US-09-018-351-3
: Sequence 3, Application US/09018351
: Patent No. 6096507
:
: GENERAL INFORMATION:
: APPLICANT: Weinshank, Richard L.
: APPLICANT: Branchek, Theresa
: APPLICANT: Hartig, Paul R.
: TITLE OF INVENTION: DNA ENCODING A HUMAN 5-HT1E RECEPTOR
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Cooper & Dunham LLP
: STREET: 1185 Avenue of the Americas
:

```

CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/018.351  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/542.358  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 1795/39317-22/JPM/MAT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-278-0400  
TELEFAX: 212-391-0525  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 422 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
US-09-018-351-3

Query Match 13.0%; Score 343.5; DB 3; Length 422;  
Best Local Similarity 22.0%; Pred. No. 4.1e-20;  
Matches 101; Conservative 92; Mismatches 190; Indels 77; Gaps 9;

26 ISLAHGIRSVLYIFLASVGN--IYLAIVLQKPKQLLOVTRFIFNLVTDLLQSL 83  
31 VTVSIVYITSLTLLTFCVAVGNACVVAATALER--SLQNVAVNLIGSLAVTDLMVSVL 88  
84 VAPWVAVTSVPLEFPLNSHFCALVSILHFAFASVNTIVLVSVDRYLSIHLPLSYPSKM 143  
89 VLPMALYQVNLKMTLGGVTCDDLFIALDVLCCTSSILHCAIALDRWAIIDPIDYVKKR 148  
144 TORRGYLLTGWTVAILOSTPPPLYGWGQAADERNALCSMIGASPSYTLISVVSFTVI 203  
149 TPRRAAALISLTWLIGFLISIPMLGWTP--EDRSPDCACTISKDHGYTISFGAFYI 206  
204 PLIYMIACYSVFCARQNALLVNKRHSLEVAVKDCVENEDEGAKEKKEPFDESFR 263  
207 PLLMLVLYGRIFRAAR-----FRIKTVKKVEKTGADTRHGASPAQPK 251  
264 ROHGEVAKAGRM--EAKDGL-----KAKEGSTGSESSVEARGSEVESSSTVASDG 316  
252 KSVNGESGRMWRIGVESKAGCALCANGAVROGDGALEIVH----- 296  
317 SMEKEGSTKVEENSMKADKGRTEVNOCSIDLGEDMEFEGDDINFSEDDVEAVNIPESL 376  
297 ---RVGNSK-EHLPLPSEAGPTPCAPASFE----- 322  
377 PPSRRNSNSNPRLRCYQCKAKAYIFIIIFSYVSLGPRYCLAVLAVWVDVETOVPQWVI 436  
323 KRNERNNAKRRMLARERKVKTLGILMGFLICWLPFFTLVAVLPCESSCHMPTLLG 382  
437 TIITLFLQCCIPHYVGYVHKTIKKEIODMLKKFPCKE 476  
383 AIIMKLGYSNLNLPVITAYTNKDPQNAFKIITICLPCRQ 422

RESULT 13

PCT-US93-00149-3  
Sequence 3, Application PC/TUS9300149  
GENERAL INFORMATION:  
APPLICANT: Weinsbank, Richard L.  
APPLICANT: Branchek, Theresa  
APPLICANT: Hartlig, Paul R.  
TITLE OF INVENTION: DNA ENCODING A HUMAN 5-HT1F RECEPTOR AND  
TITLE OF INVENTION: USES THEREOF  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/00149  
FILING DATE: 19930108  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 1795/39318  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-977-9550  
TELEFAX: 212-664-0525  
TELEX: 422523 COOP UI  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 422 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: unknown  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
IMMEDIATE SOURCE:  
CLONE: 5-HT1A  
PCT-US93-00149-3

Query Match 13.0%; Score 343.5; DB 5; Length 422;  
Best Local Similarity 22.0%; Pred. No. 4.1e-20;  
Matches 101; Conservative 92; Mismatches 190; Indels 77; Gaps 9;

26 ISLAHGIRSVLYIFLASVGN--IYLAIVLQKPKQLLOVTRFIFNLVTDLLQSL 83  
31 VTVSIVYITSLTLLTFCVAVGNACVVAATALER--SLQNVAVNLIGSLAVTDLMVSVL 88  
84 VAPWVAVTSVPLEFPLNSHFCALVSILHFAFASVNTIVLVSVDRYLSIHLPLSYPSKM 143  
89 VLPMALYQVNLKMTLGGVTCDDLFIALDVLCCTSSILHCAIALDRWAIIDPIDYVKKR 148  
144 TORRGYLLTGWTVAILOSTPPPLYGWGQAADERNALCSMIGASPSYTLISVVSFTVI 203  
149 TPRRAAALISLTWLIGFLISIPMLGWTP--EDRSPDCACTISKDHGYTISFGAFYI 206  
204 PLIYMIACYSVFCARQNALLVNKRHSLEVAVKDCVENEDEGAKEKKEPFDESFR 263  
207 PLLMLVLYGRIFRAAR-----FRIKTVKKVEKTGADTRHGASPAQPK 251  
264 ROHGEVAKAGRM--EAKDGL-----KAKEGSTGSESSVEARGSEVESSSTVASDG 316  
252 KSVNGESGRMWRIGVESKAGCALCANGAVROGDGALEIVH----- 296  
317 SMEKEGSTKVEENSMKADKGRTEVNOCSIDLGEDMEFEGDDINFSEDDVEAVNIPESL 376

Db 297 ---RVGNSK-EHLPSEAGPTPCADAPSE----- 322  
Qy 377 PPSRNSNSNPPLPRCQCQAKAVIFIIIFSVLSLGPYCFPLAVLAWVDVETQYPMVI 436  
Db 323 KKNERNNAKRMALAREKVKTLGIITGFTLCLWLPFTFIVLVLPCSSCHMPTLLG 382  
Qy 437 TIIMLFLOCCIHPIYVGYMHKTIKEIDMLKFFCKE 476  
Db 383 AIIWLGTSNLSLNPVIYAFNKDFQNAFKKIICLCFCRQ 422

RESULT 14  
US-08-334-698-6  
; Sequence 6, Application US/08334698  
; Patent No. 5556753  
; GENERAL INFORMATION:  
; APPLICANT: Jonathan A. Bard et al.  
; TITLE OF INVENTION: DNA Encoding Human Alpha 1 Adrenergic  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: COOPER & DUNHAM  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.24  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/334,698  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/952,798  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 376901  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 977-9550  
; TELEFAX: (212) 664-0525  
; TELEX: (212) 422523 COOP UI  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 466 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-334-698-6

Query Match 12.8%; Score 338.5; DB 1; Length 466;  
Best Local Similarity 21.3%; Pred. No. 1.2e-19;  
Matches 108; Conservative 80; Mismatches 187; Indels 131; Gaps 11;

Qy 8 STRESNSSHTCMPLSKMPLSLAHGIIRSTVLVIFLAASFGVNIYALVLOKRPOLQYTN 67  
Db 5 SGNASDSSNCRQP--PAVNVISKAILLGVILGILFVGLNLIYILSVACHRHLSVTH 62  
Qy 68 RFIINLVTDLLQISLVAAPVVAATSVPLFWPLNSHCTALVSLTHLFAFASVNTIVLSV 127  
Db 63 YYIYNLAVADLLITSTVLPFSAIFEVGLGYMAFGVFCNIMAVADVLCCTASIMGLCIISI 122  
Qy 128 DRYLSIIHPLSPSKMQRGRGYLLLYGWIYAILLOSTPPLYGWGAADDERNALCSMTWG 187  
Db 123 DRYGVSPLEKPYITVYQRGIMALLCWMALSLVISISPLFGWRQPA--PEDETTICQI--N 179  
Qy 188 ASPSTIISVSVFIYPLIVIAICYSVFCARROHALLVNWKRHSLEVRKDCVENDE 247

Db 180 EEPGYLFSALGSFYPLPLAILVMYCRVYVAKR-----ESRGL 218  
Qy 248 EGAEKKEFQDESEFRROHBEVAKAKERMAMKOSLKAKEGSGTSSVEARSGSEYR 307  
Db 219 KSGLTKDSDSEQYTLRIHRNAPR-----GGSGMASAKTKTHFSVRL- 261  
Qy 308 ESSVVASDGSMEKEGSGTKVEENSMKADKGRTEVNOCSIDGEDMGEEDINFESEDV 367  
Db 262 -----LFSRER- 268

Qy 368 EAVNIPESLPSRNSNSNPPLPRCQCQAKAVIFIIIFSVLSLGPYCFPLAVLAWVDV 427  
Db 269 -----KAKTLGIVGCVFLCWLPE-FLVMPIGSFP 299  
Qy 428 ETQYPMVITIIIMLFLOCCIHPIYVGYMHKTIKEIDMLK-KFCKEKPPE-----D 482  
Db 300 DKFSEIVFKIVFWGLYNSCINPLIYVCSQEFKAPQNLRIQCLCRKOSKHALGT 359

Qy 483 SHPLPGTEGTEGKI-VPSYDSATF 507  
Db 360 LHPSSQAVEGQKDMVRIPVGSRETF 385

RESULT 15  
US-08-228-932-6  
; Sequence 6, Application US/08228932  
; Patent No. 5578611  
; GENERAL INFORMATION:  
; APPLICANT: Charles Gluchowski, Carlos C. Forray, George Chiu,  
; APPLICANT: Theresa A. Branchek, John M. Wetzel and Paul R. Hartig  
; TITLE OF INVENTION: USE OF ALPHA-1C SPECIFIC COMPOUNDS TO TREAT BENIGN  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: COOPER & DUNHAM  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.24  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/228,932  
; FILING DATE: 13-APR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 41878-B/JPM/TTP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 977-9550  
; TELEFAX: (212) 664-0525  
; TELEX: (212) 422523 COOP UI  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 466 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-228-932-6

Query Match 12.8%; Score 338.5; DB 1; Length 466;  
Best Local Similarity 21.3%; Pred. No. 1.2e-19;  
Matches 108; Conservative 80; Mismatches 187; Indels 131; Gaps 11;

Qy 8 STRESNSSHTCMPLSKMPLSLAHGIIRSTVLVIFLAASFGVNIYALVLOKRPOLQYTN 67  
Db 5 SGNASDSSNCRQP--PAVNVISKAILLGVILGILFVGLNLIYILSVACHRHLSVTH 62

```

QY 68 REFINLVTDLQSLVAAPVAVASVPLFAMNSHFTALSTVLSHFAFASVMTVIVSV 127
Db 63 YIYNLVAADLLLTSTVLPFAIEVGLGMAFGVFCFNMAADVLCCTASINGLCTISI 1222
QY 128 DRYLSIHPLSYPSKMTQRRGYLLLTSTVVAILOSTPRPLYGMOQAEDERNALCSMIWG 1877
Db 123 DRYIGVSFPLRYPTLIVRRGLMALLCVMALSTVLSIGPLFGMRQPA - PEDETICQI - N 179
QY 188 ASPETILTSVSFLIYPLIPLVIMACYSVFPCARRQOHLNLVNRKHSLEVAVKCOVENEDE 2477
Db 180 EEPGVLFSAIGSYLLPLATLILVMTCRYVYAKR-----BSRGL 218
QY 248 EGAEKKEEFODESEFRRQHEGEVNAKEGMRBKAQSLKAKEGSYGTSESSVEARGSEVR 3077
Db 219 KSGLKTKQKSDSEQYTLRIHRKNAPA-----GSGGMSAKTKTHFSYRL- 261
QY 308 ESSIVASDGSMEKGEGSTKYEENSMKADKGKTEYNOCSIDLGEDMFEDEDINFSDDV 3677
Db 262 -----LKFSREK- 268
QY 368 EAVNIPESLPSRRNSNSNPPLPRCYQCKAKAVFIIFISYVLSIGPYCFVLAVLAVADV 4277
Db 269 -----KAKTIGIYVGCFCVLQWLPF -FLVMP IGSFEP 2999
QY 428 ETQVQWVITIIILWLFLOCCIHBYVYGVMHKTIKKEIDMLK -KFECKEKPRKE---D 4828
Db 300 DFKSESEYFEKIVFWLGYLNSCINFLIYPCSSQEFKKAQVNLRIQCLCRROSSHMLGYT 3599
QY 483 SHPDLPGTEGTECKI -VPSIDSKTF 507
Db 360 LHPSQAVEGQHKDMVRIPIPSRETF 385

```

Search completed: October 28, 2002, 16:47:28  
Job time : 19 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 16:44:33 ; Search time 21 Seconds  
(without alignments)  
2324.447 Million cell updates/sec

Title: US-09-992-238-2

Perfect score: 2644

Sequence: 1 MTSTCTNSPRESNSHTCMP.....GTGGTGKIVPSYDSATPP 508

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	343.5	13.0	490	2 A35546	muscarinic acetylcholine receptor M4 - chicken
2	340.5	12.9	422	2 I38209	serotonin receptor
3	340	12.9	477	2 S71323	alpha-1A adrenergic
4	336	12.7	466	2 I57959	alpha-1C adrenergic
5	335.5	12.7	421	2 I49375	serotonin receptor
6	333.5	12.6	429	2 S65656	alpha-1C adrenergic
7	333.5	12.6	466	2 JN0765	alpha-1C adrenergic
8	333.5	12.6	499	2 S65657	alpha-1C adrenergic
9	331.5	12.5	422	2 JH0315	serotonin receptor
10	325.5	12.3	466	2 A35375	alpha-1 adrenergic
11	319	12.1	509	2 A47174	serotonin receptor
12	314.5	11.9	501	2 JH0447	alpha-1A adrenergic
13	314.5	11.9	572	2 I39369	alpha-1A adrenergic
14	312.5	11.8	564	2 S48657	muscarinic acetylcholine receptor
15	306.5	11.6	460	2 A38731	muscarinic acetylcholine receptor
16	304.5	11.5	601	2 JH0170	octopamine receptor
17	296.5	11.2	601	2 S12004	tyramine receptor
18	294	11.1	466	2 S10856	muscarinic acetylcholine receptor
19	291	11.0	466	2 JH0197	muscarinic acetylcholine receptor
20	291	11.0	479	2 S10127	muscarinic acetylcholine receptor
21	289	10.9	466	2 A27386	muscarinic acetylcholine receptor
22	288	10.9	488	2 T15941	hypothetical protein
23	288	10.9	515	2 JCI525	hypothetical protein
24	287	10.9	466	2 S10126	muscarinic acetylcholine receptor
25	285	10.8	610	2 T32917	hypothetical protein
26	283	10.7	515	2 A40491	alpha-1 adrenergic
27	283	10.7	517	2 A45121	alpha-1B adrenergic
28	282.5	10.7	478	2 C29514	muscarinic acetylcholine receptor
29	282	10.7	479	2 S33776	muscarinic acetylcholine receptor

30	280	10.6	466	2 A40972	muscarinic acetylcholine receptor
31	279.5	10.6	476	2 JC5042	G protein-coupled
32	277.5	10.5	531	2 JT0531	muscarinic acetylcholine receptor
33	276	10.4	564	2 A38271	serotonin receptor
34	275.5	10.4	484	2 S58868	G protein-coupled
35	275.5	10.4	532	2 JT0530	muscarinic acetylcholine receptor
36	273	10.3	443	1 DYHND2	dopamine receptor
37	273	10.3	445	2 A48881	serotonin receptor
38	270	10.2	448	2 A47519	serotonin receptor
39	268	10.1	448	2 S36402	serotonin receptor
40	267.5	10.1	442	1 DYXLD2	dopamine receptor
41	265	10.0	501	2 T18863	hypothetical protein
42	262.5	9.9	445	2 T42203	serotonin receptor
43	260	9.8	430	2 T16079	hypothetical protein
44	259.5	9.8	589	2 B29514	hypothetical protein
45	258	9.8	460	2 A31897	muscarinic acetylcholine receptor

ALIGNMENTS

RESULT 1

A35546 muscarinic acetylcholine receptor M4 - chicken  
C:Species: Gallus gallus (chicken)  
C>Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 21-Jul-2000  
C:Accession: A35546  
R:Titelje, K.M.; Goldman, P.S.; Nathanson, N.M.  
J. Biol. Chem. 265, 2828-2834, 1990  
A:Title: Cloning and functional analysis of a gene encoding a novel muscarinic acetylcholine receptor  
A:Reference number: A35546; MUID:90153912  
A:Accession: A35546  
A:Molecule type: DNA  
A:Residues: 1-490 <TIE>  
A:Cross-references: GB:J05218; NID:g211067; PIDN:AAA48563.1; PID:g211068  
C:Superfamily: vertebrate rhodopsin  
C:Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; transmembrane protein; status predicted <TM1>  
F:42-67/Domain: transmembrane #status predicted <TM1>  
F:80-104/Domain: transmembrane #status predicted <TM2>  
F:117-138/Domain: transmembrane #status predicted <TM3>  
F:158-181/Domain: transmembrane #status predicted <TM4>  
F:203-226/Domain: transmembrane #status predicted <TM5>  
F:413-433/Domain: transmembrane #status predicted <TM6>  
F:445-466/Domain: transmembrane #status predicted <TM7>

Query Match 13.0%; Score 343.5; DB 2; Length 490;  
Best Local Similarity 23.7%; Pred. No. 8.2e-16;  
Matches 116; Conservative 94; Mismatches 176; Indels 103; Gaps 17;

QY	36	TVLVIFLA-----ASFVGNIVLALVLRKPKQLQVTFNFINLVTDL-----QISL 83
DB	39	TVELVFATVGTGSLVTVVGNLILVMSIKVNRQLQVNNFELSLACADILIGVFSMNL 98
QY	84	VAPVWVATSVPLFWPLNSHFTALVSLTLFAPASVTIVTVSVDRSLSIHPSPSKM 143
DB	99	YTVYIING----YMPGLAVVCDLWLDLDYVSNVSNMILLISDRFCVYKPLFYARR 154
QY	144	TORRGYLLVGTWITVAIILOSTPPLYGW---GQAAPFERNALCSMIGASPSYITLSVS 199
DB	155	TTKAGLMIAAAWILSLFLMAPALIFWQFIWKRRTVHERE--CYIQFLSNPAVFGFAIA 212
QY	200	FIVPLVIMACYSVVCAARQHALLVNKRHSLEVR-----VKDEVN 244
DB	213	AFYPLVIMVLYLHIISLASR-----VRRHPESSRERKSKSLFFKAPVKKNNN 266
QY	245	EDEGAEKKEEFODESEFRROHEGEVAKKEGRMAKGSILKAKGSTGTSSEVSARGSE 304
DB	267	SPKAAVEKKEEVRN-----GKV-----DDPSAQCEAIGQGE-----KE 301
QY	305	EVRESSIVASDGSMEGKSGTRVEENSMKADKGRTEVNGSIDLEGDMERGEEDINFESE 364
DB	302	TSNESSIVSMQTQTKDRP-TTEILPAGOGGSPAPRVPPTS-----KMSKIKIYTKQ 352

```

OY 365 DDVEANIPESLP-----PSRRN--SNSNP-LPRCY-----OCCAAKY 400
DB 353 TGTSESTATETIVPKACASDHNSLSNSRPANVARKKFAISAOVKKKROMAAREKKVTTRI 412
OY 401 IFTIFSYVSLGCPCEFLAVLAVWVDVETQVPQVYITIIIMFLPQCCIHPPVYGYMKT 460
DB 413 IFAILLAFILITWIPYNNVNLINFE--CETCVPEIYWSIGYWLICYNSTINPACYALCNMT 470
OY 461 IKKEIDML 469
DB 471 FKKTFFKHL 479

RESULT 2
138209
N:serotonin receptor 1A - human
N:Alternate names: 5-hydroxytryptamine receptor 1A (5-HT1A)
C:Species: Homo sapiens (man)
C:Date: 06-SEP-1996 #sequence.revision 15-May-1998 #text_change 21-Jul-2000
C:Accession: I38209; I37104; S07343; I56176; S31438
C:Jacobsen, J.S.; Refolo, L.M.; Conley, M.P.; Sambamurti, K.; Humayun, M.Z.
A>Title: DNA replication-blocking properties of adducts formed by aflatoxin B1-2,3-dichloroacetate. Res. 179, 89-101, 1987
A:Reference number: I38209; MUID:87258013
A:Accession: I38209
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-422 <AC>
A:Cross-references: EMBL:X57829; NID:g36428; PIDN:CAA040962.1; PID:g36429
R:Parkes, C.L.; Chang, L.S.; Shenk, T.
Nucleic Acids Res. 19, 7155-7160, 1991
A>Title: A polymerase chain reaction mediated by a single primer: cloning of genomic sequences. Nature 329, 75-79, 1987
A:Reference number: I37104; MUID:92115564
A:Accession: I37104
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-9 <PAR>
A:Cross-references: EMBL:Z11168; NID:g1033027; PIDN:CAA77560.1; PID:g1033028
R:Koblika, B.K.; Fritelle, T.; Collins, S.; Yang-Feng, T.; Koblika, T.S.; Francke, U.; Le
Nature 329, 75-79, 1987
A>Title: An intronless gene encoding a potential member of the family of receptors coupled to G proteins. Nature 329, 75-79, 1987
A:Reference number: S07343; MUID:87315369
A:Accession: S07343
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-151, 'R', 155-171, 'I', 173-417, 'N', 419-422 <KOB>
A:Cross-references: EMBL:X13556; NID:g35523; PIDN:CAA31908.1; PID:g35524
A:Note: the authors translated the codon ATC for residue 172 as Met
R:Aune, T.M.; McStrath, K.M.; Sarr, T.; Bombara, M.P.; Kelley, K.A.
J. Immunol. 151, 1175-1183, 1993
A>Title: Expression of 5HT1A receptors on activated human T cells. Regulation of cyclic AMP formation by 5HT1A receptors. J. Biol. Chem. 268, 11111-11116, 1993
A:Reference number: I56176; MUID:93329096
A:Accession: I56176
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: RPR, 203-227, 'R', 229-244, 'A', 245-354, 'T', 356-362, 'MRP' <AUN>
A:Cross-references: GB:S64045; NID:g404416; PIDN:AD13945.1; PID:g4261645
C:Genetics:
A:Gene: GDB:HTRLA; ADRB2RL1; ADRBR1L
A:Cross-references: GDB:120686; OMIM:109760
A:Map position: Scen-5q11
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; lipoprotein; neurotransmitter receptor
F:37-62/Domain: transmembrane #status predicted <TM1>
F:74-98/Domain: transmembrane #status predicted <TM2>
F:110-133/Domain: transmembrane #status predicted <TM3>
F:153-178/Domain: transmembrane #status predicted <TM4>
F:192-211/Domain: transmembrane #status predicted <TM5>
F:346-367/Domain: transmembrane #status predicted <TM6>
F:379-403/Domain: transmembrane #status predicted <TM7>
F:109-187/Disulfide bonds: #status predicted
F:420/Binding site: palmitate (Cys) (covalent) #status predicted

```

```

Query Match          12.9%: Score 340.5; DB 2; Length 422;
Best Local Similarity 22.0%: Pred. No. 1,1e-15;
Matches 101; Conservative 92; Mismatches 190; Indels 77; Gaps 9;

QY      26 ISLHGIIIRSTVLVYFIPLAASEFVGN--IVLALVLRKRPOLLQVTRNFIENLLVTDLQISL 83
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      31 VTSYQVITSLTLGLTGLFCVAILGNACVVAALALER--SLQVNAVYLIGSLAVTDLMTSVL 88
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      84 VAPWVATSVPLFPMPLNSHFCTALVSLTHLFAFASVNTIYLVSDRLSTIHPLSYSKM 143
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      89 VLPMAALYQVNLKMTPLGGVTCDFLALDVLCCSTSIHLCAIALDRWALTDPIDYVNRK 148
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      144 TORRGYLLTYGTWYIAIIQSTPPLYGMGAAPFERNALCSTMIGASPSYTLISVSEFVI 203
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      149 TPRRAAALISLTWLIGFLISIPMLGWTRP--EDRSDDACTISKDHQYITYSIFGAFYI 206
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      204 PLIVMIACYSVFPCARROHALLVNRKHSLEVRKDCVENEDDEGAEKKKEFODESEFR 263
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      207 PLTLMLVLYGRIFRFAAR-----FRIKRYKKKVEKTEGADTRHGASAPAPQK 251
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      264 RQHEDEVAKKEGRM--EAKDQSL-----KAKEGSTGTSSESVKRGSEVRESSTVASDG 316
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      252 KSVNSESSESRMRWRIGVESKAGCALCANGAVRQGDGALVEYEH----- 296
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      317 SMEKEGSTKYEENSMKADKGRTEVNOCSIDLGEDDMFEGEDDINFSEDDVEAVNIPESL 376
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      297 ----RVGNSK -EHLPLPSEAGTPPCAPASFE----- 322
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      377 PPSRRNSNSNPPLPRCYOCKAAKAYIFIIISYVLSLGRPYCLAVLAWVDYETVPQWVI 436
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      323 RKNENAAEAKKRMALARERKTYVKTIGITGTFIIICWLPFFIYALVLPFCSESSCHMPTLLG 382
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      427 TTIIMLFLOCCIHPRVYGVYMKTKIKKEIDMKKPFCKE 476
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      383 AITNWLGSNSLNPVITAYFENKDFQNAFKTIKCKPFRQ 422

RESULT 3
S71323
alpha-1A adrenergic receptor - Japanese medaka
C:Species: Oryzias latipes (Japanese medaka)
C:Date: 23-Jul-1997 #sequence_revision 01-Aug-1997 #text_change 17-Mar-1999
C:Accession: S71323
R:Yasuoka, A.; Abe, K.; Arai, S.; Emori, Y.
Eur. J. Biochem. 235, 501-507, 1996
A>Title: Molecular cloning and functional expression of the alpha-1-A-adrenoceptor of
A:Reference number: S71323; MUID:96184522
A:Accession: S71323
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-477 <YAS>
A:Cross-references: EMBL:D63859
A>Note: It is uncertain whether Met-1 or Met-8 is the initiator
C:Superfamily: vertebrate rhodopsin

Query Match          12.9%: Score 340; DB 2; Length 477;
Best Local Similarity 23.8%: Pred. No. 1,4e-15;
Matches 119; Conservative 66; Mismatches 180; Indels 134; Gaps 14;

QY      12 SNNSTCMPPLSKMPTISLHGIIIRSTVLVYFIPLAASEFVGNIVLALVLRKRPOLLQVTRNFI 71
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      18 SMCSSVLAIP---ELNTKAVAVLGVLTGIFLFGVGNILVILSVCHRHQDTYYTYFI 73
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      72 NLVYTDLQISVAPWVATSVPLFPMPLNSHPCATVLSLTHLFAFASVNTIYLVSDRYL 131
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      74 NLAADVADLLSLSTVLPFSALFETLDRWVFGVPCNIMAAVADVLCCTASIMSLCVISVDYRI 133
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      132 STIHPILSPKMTORRGYLLTYGTWYIAIIQSTPPLYGMGAAPFERNALCSTMIGASPS 191
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      134 GVSYFLRPAITKTRRALLAWMLVLSVIISIGLFPWKRPKPA -PEDETVCYKIT -DEPG 190
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      192 YTIISVSEFIVPLIVMIACYSVFPCARROHALLVNRKHSLEVRKDCVENEDDEGAE 251

```

Query Match	12.7%	Score 356	DB 2	Length 466
Best Local Similarity	21.3%	Pred. No.	2.5e-15	
Matches 108	Conservative 79	Mismatches 166	Indels 134	Gaps 11
QY	7	NSTRSSNSHHCMPLSKMPISLHGIIRSYLVIFLAASFQNVITAIYLRKPOLLOVY	66	
		,   :::	:	:   ::

Query Match	12.7%	Score 335.5;	DB 2;	Length 421;
F:345-36//Domain: transmembrane	#status predicted <TM>			
F:379-403//Domain: transmembrane	#status predicted <TM>			
F:10_11,24,30//Binding site: carbohydrate (Asn) (covalent)	#status predicted			
F:109-18//Disulfide Bonds:	#status predicted			





A:Accession: A35181  
A:Molecule type: DNA  
A:Residues: 1-372, 'S', 374-422 <ALB>  
A:Cross-references: GB:J05276; NID:g202540; PIDN:AAA0612.1; PID:g202541  
C:Superfamily: vertebrate rhodopsin  
C:Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; transmembrane  
F:37-62/Domain: transmembrane #status predicted <TM1>  
F:74-98/Domain: transmembrane #status predicted <TM2>  
F:109-134/Domain: transmembrane #status predicted <TM3>  
F:153-178/Domain: transmembrane #status predicted <TM4>  
F:252-277/Domain: transmembrane #status predicted <TM5>  
F:346-371/Domain: transmembrane #status predicted <TM6>  
F:379-404/Domain: transmembrane #status predicted <TM7>

Query Match	12.5%	Score 331.5	DB 2	Length 422
Best Local Similarity	23.5%	Pred. No. 4.5e-15		
Matches 11; Conservative	76	Mismatches 181	Indels 105	Gaps 12

QY	26	ISLGGIIHSYLVLFVFLAASFPVGN--IYALVQJRKPOLQVYTNFIFLTLDTLLQSL	83
b	31	VYFSTQVITVSLLGTLIFCAVGNACVVAATALER--SLQNVANTLIGSLAVTDLMVSVL	88
QY	84	VAPVVVATSVPLFWPLNSHFTALVSLTHLFAFASVNTIVLVSVDYRYSIIHPLSPSKM	143
Db	89	VLPMAALYQVLKMWLTIGQYTCDFLADVLDTCTSSLIHCAIALDRYMAITPDIYVNMK	148
QY	144	TORGITLLGYWIAVAILTOSTPPLLYGWGOAAMDERNALCSTMIGASPSPTIIISVFIYI	203
Db	149	TFRRRAAALISLTMWLTIGFLISIPRMGWRP--EDRDPDACTISKDHGYTIISTGCIFYI	206
QY	204	PLIVMIACISVYFCARQOHALLVNYKRSLSEVRVACDVENDEDEGAKEEFQDESEFR	263
Db	207	PLILMLVLYGRIFRAAR-----FRIRKRYKVEKKGA-----	238
QY	264	ROHEGEVKAKEGRMEAKDGSLKAKBGSTGTSESSVSEARGSEEVREBSSTIVASDQSMGKEG	323
Db	239	-----GTS-----LGTSSAPPRKSLINGQPG	259
QY	324	S---TKVEENSMKADKGRTEVNOCSIDLQDDEMERGEDDINFSEDDVEAV-NIPESLP-P	378
Db	260	SGDMRCAEN-----RAVGTPCT-----NGAVRQDDDEATLEVTIEVHRVNGSKHEHLP	308
QY	379	SRRNSNSNRP-----LPRCYOCKAKAVFIIFIIISVYLSLGPYCFPLAVLA	422
Db	309	SSGSGNSYAPACLERKNEKNEAERKRMALAREKTYKVTGLIIMGPFIILCMLPFIYALVL	368
QY	423	VWVDVETQVPOWYITIIIMVLFLOCCIHPRVYGYMHKTIKTKELOMLKPFCK	475
Db	369	PCPCENSHAPBALGAIIMWLGYSNSLNPVIAYFPRKQDQNAFKILKKCFGR	421

RESULT 10  
A53375  
alpha-1-adrenergic receptor - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 13-Aug-1999  
C:Accession: A53375  
R:Schlamm, D.A.; Lomasney, J.W.; Lorenz, W.; Szklut, P.J.; Fremean Jr., R.T.; Yang-Feng  
J. Biol. Chem. 265, 8183-8189, 1990  
A:Title: Molecular cloning and expression of the cDNA for a novel alpha-1-adrenergic re  
A:Reference number: A53375; MUID:90243698  
A:Accession: A53375  
A:Molecule type: mRNA  
A:Residues: 1466 <SC>  
A:Cross-References: GB:J05426; NID:g162663; PIDN:AA30374.1; PID:g162664  
C:Superfamily: vertebrate rhodopsin  
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane pro

Query Match 12.3%; Score 325.5; DB 2; Length 466;  
Best Local Similarity 20.4%; Pred. No. 1.3e-14;  
Matches 107; Conservative 78; Mismatches 182; Indels 157; Gaps 10;

7 NSTRESNSSHTCMLSKMPI SLAHGIRSTVLVIFLAASFVGNIVLALVLQRPQLQVT 66

```

Db 7 NASDSNCTH-----PPPPVNIKAILLGGVILGGLLFGVILGNIIVLITLVSVAOCHRLHSVT 61
QY 67 NREIFNLLVTDLLQISLVAPVVATSVPLEMPLNSHCTALVSLTHFAFASVNTIVLVS 126
Db 62 HYYINLVAADLLLTSTVLPFSALFEILGYNAFGVFCFNWAAADVLOCCTASIMGLCTIS 121
QY 127 VDRYLSTIHLPLSPKMNORRGYLLLTGTWYVALIQLSTPRPLVGOQAAFDERNALCSIMW 186
Db 122 IDRYGVSTPLKPTPIVYQKRGMLALCVMALSTLVSIGPLFGWKQPA-PEDETICQI-- 178
QY 187 GASPSYTLISVVSFIVPIVIMACYSVFCARQOHALLVYKRNHSLVBRKYDCVENED 246
Db 179 NEEPGVYLFSAIGSTYVPLTILVMKCRYVYAKR----- 213
QY 247 EEGAKEKEEFODESEFRROHEGEVAKAKEGRMEAKDGLKAKEGSTGTSESSVEARGSEEV 306
Db 214 -----ESRG----- 217
QY 307 RESSTVASDGSMEKEGSKTVEENSMKADKGTVEYNQCSIDLEBDMFEGEDDIFSEDD 366
Db 218 -----LKSGLTKTDSDS---EQVTLRIHRKNAOYGGSGSVSAKKTIHRYVLRKFSREK 268
QY 367 VEAVNIPESLPPSRNNSNPNPLPRCYOCCAKAKVFIIFSVLSLGPYCFALAVLAWVD 426
Db 269 -----KAAKTGLIYGCVCYLQWLPF-FLVMPIGSFF 298
QY 427 VETQVPQWVITIIIMFLFLOCCIHRYVYGYMHKTIKKELIOMLKKKFFCKRPPK----- 480
Db 299 PPFRRSEYVFKIAFMGLYINCINPIIYPCSSQEFKKAFQVNLRIQCLRRKQSKHTLGY 358
QY 481 -----EDSHPL-----PTEGTEBEKIYPS 501
Db 359 TLHAFSHVLEGOHKDLVRLPVGSAETFYKISKTDVCEMKRFS 402

```

RESULT 11

AA7174  
serotonin receptor - great pond snail  
N:Alternate names: 5-hydroxytryptamine receptor (5-HT<sub>2</sub>)  
C:Species: *Lymnaea stagnalis* (great pond snail)  
C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 24-Sep-1999  
C:Accession: AA7174  
R:Suzumori, K.S.; Sunahara, R.K.; Guan, H.C.; Bullock, A.G.; Tensen, C.P.; Seeman, P.  
Proc. Natl. Acad. Sci. U.S.A. 11-15, 1993  
A:Title: Serotonin receptor cDNA cloned from *Lymnaea stagnalis*.  
A:Reference number: AA7174; MUID:93126323  
A:Accession: AA7174  
A:Status: preliminary  
A:Molecule type:nucleic acid  
A:Residues: 1-509 <SDS>  
A:Cross-references: GB:LO6803; NID:q159456; PIDN:AAA29290.1; PID:q159459  
A:Experimental source: central nervous system  
A>Note: sequence extracted from NCBI Backbone (NCBIN:121945, NCBIF:121952)  
C:Superfamily: octopamine receptor type I  
C:Keywords: neurotransmitter receptor

[illegible]

QY	3	SCHTCSFRESNSHRCMPKSPKBPISLAHHIIRSTVAVIFLASFPGN--IVATLVQLORP	60
	:	: : : : : : : : : : : : : : : :	
Dd	74	TGATNTDTRMSLY-----YSHEHLVLTSLGLFLVLCICGNCFVIAAMLER--	124
QY	61	QLLOQTNRRTFNLLVDLIQISLVAPMVYAITSVPLEPLANSFCALVSITHLFAFSYN	120
	:	: : : : : : : : : : : : : : : :	
Dd	125	SHINAAVYILISLAADVAALVAMPLSVASESKWPFJHSEYCKDMISDYDLCTASTL	184
QY	121	TTLVSVDRYSIIHPLSYPSKMTORGYLELLXGTWIAVAILDSTPEPLYGSGAARD-ERN	179
	:	: : : : : : : : : : : : : : : :	
Dd	185	HLLVALAMDHYMAVT-SIDVIRRSARRILLIMVMVIALFTSIPLEFWGRPNDDPKT	243



```

QY 268 GEVAKK3E9RMAK0GSLKAKEGSTJSSSEVAKGSEVRRSSVADGSEKGEKSTKV 327
    || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 290 AGVRRERG-----KASEYVLR--HCRGAATGADGA-- 318
QY 328 EENSKAKDKGRTENVQCSIDLGEDMEFEGEDDINFSEDDVEAVNIPESLPPRSNSNP 387
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 319 -HGMRSAKGHTFRNSLSVRL--LAFSRK----- 344
QY 368 PLPRCYCKAKAVFIIFISVYLSIGPCFLAVLAAMVDVETQVPPWVITIIIMFLQOC 447
    |||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 345 -----KAAKTIAIVGVFLCWFMPFFFLPKGS-LFPQLKPSGEVFKVIEFWLGYENS 395
QY 448 CIHPYVVGMMHTIKELQDMLK 470
Db 396 CVNPLIYPCSSSEERKAFRLRLR 418

```

## RESULT 14

muscarinic acetylcholine receptor MR - African clawed frog  
 :Species: *Xenopus laevis* (African clawed frog)  
 :Date: 15-Jul-1995 #sequence\_revision 10-Nov-1995 #text\_change 21-Jul-2000  
 C:Accession: S48657; S24948  
 R:Herrera, L.; Carvallo, P.; Antonelli, M.; Olate, J.  
 FEBS Lett. 352, 175-179, 1994  
 A:Title: Cloning of a *Xenopus laevis* muscarinic receptor encoded by an intronless gene.  
 A:Reference number: S48657; MUID:95010703  
 A:Accession: S48657  
 A:Molecule type: mRNA  
 A:Residues: 1-484 <HER>  
 A:Cross-References: GB:X65865; NID:g64900; PIDN:CAA6694.1; PID:g64901  
 R:Olate, J.  
 submitted to the EMBL Data Library, April 1992  
 A:Reference number: S24948  
 A:Accession: S24948  
 A:Molecule type: mRNA  
 A:Residues: 1-131, 'X', 133-484 <OLA>  
 A:Cross-References: EMBL:X65865; NID:g64900; PID:g64901  
 C:Superfamily: vertebrate rhodopsin  
 C:Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; transmembrane  
 F:33-58/Domain: transmembrane #status predicted <TM1>  
 F:71-95/Domain: transmembrane #status predicted <TM2>  
 F:108-129/Domain: transmembrane #status predicted <TM3>  
 F:149-172/Domain: transmembrane #status predicted <TM4>  
 F:194-217/Domain: transmembrane #status predicted <TM5>  
 F:407-427/Domain: transmembrane #status predicted <TM6>  
 F:439-460/Domain: transmembrane #status predicted <TM7>

Query Match	11.88; Score 312.5; DB 2; Length 484;
-------------	---------------------------------------

Best Local Similarity 21.0%; Pred. No. 1e-13;

Matches 110; Conservative 105; Mismatches 209; Indels 99; Gaps 16;

7 NSTRESNSSHTCMLSKMPISLAHGIRSTVLVIFLA-----ASFVGNIVLALVLQR

3 NDIWENESSASNSHIDETIVEIPGRY--QTMEMIFIAIVTGSLSLTVVGNILVMSIKV

39 KPOLLQVTKRETFNLEVIDL-----QISLVAPWVAISVPLEFWPLNSHCIALVSLIHLF 114

01 NKQLQI VNNI ELSLACADLLIGVFSMNLISLI LKNG --- I MPELGI VCDLWLAELDI V LI

113 ALKSVN11VLEVSVDKRL2S1LHPS1P3NM1XNRG1EELIG1W1VAL2VS1FFLIGWQAA 1/4

[illegible][illegible][illegible]

291 TGTSSSVFARGSEFVRESSTVASDGS-----EGKEGSTKVEENSMK--ADKG 337

```

Db      274  NCKT|K|S|MT|N|L|O|T|A|E|K|E|T|S|N|E|S|S|A|S|I|N|H|P|P|E|K|O|P|L|S|E|S|S|V|L|A|P|Q|S|N|P|L|P|A|K|A|
Qy      338  R|T|E|-----V|N|O|C|I|D|E|D|E|D|E|G|E|D|I|N|F|S|E|D|D|E|A|V|N|I|P|E|
Db      334  N|T|A|S|K|S|K|I|K|I|Y|K|O|T|G|N|E|C|Y|A|T|E|I|Y|E|C|A|I|P|E|-----Q|A|N|N|P|V|V|A|R|
Qy      375  S|L|P|R|R|N|S|N|P|L|P|R|C|Y|O|C|A|K|A|V|I|I|F|S|V|L|S|G|P|C|F|L|A|V|A|V|D|E|T|O|V|P|O|W|
Db      361  K|F|A|S|A|R|Q|O|V|R|K|K|R|O|A|R|E|K|V|T|R|I|F|A|I|L|A|I|T|I|T|P|N|V|V|L|N|T|F|---C|O|T|C|I|P|E|T|
Qy      435  V|T|I|I|W|E|F|L|O|C|I|H|P|V|Y|G|M|H|T|I|K|E|I|O|D|I|K|K|F|E|C|K|
Db      439  I|W|I|G|I|W|I|C|Y|N|S|T|I|N|P|A|C|Y|L|A|C|N|A|T|E|K|K|T|K|H|L|---M|C|O|Y|K|

```

## RESULT 15

alpha-1A adrenergic receptor - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 20-Apr-2000  
 A:Accession: A38731; A53280  
 R:Idonastrey, J.W.; Cotecchia, S.; Lorenz, W.; Leung, W.Y.; Schwinn, D.A.; Yang-Feng, T.  
 J. Biol. Chem. 266, 6365-6369, 1991  
 A>Title: Molecular cloning and expression of the cDNA for the alpha-1A-adrenergic rec  
 A:Reference number: A38731; MUID:91177889  
 A:Accession: A38731  
 A:Molecule type: mRNA  
 A:Residues: 1-560 <LOW>  
 A:Cross-references: GB:M0654; NID:q202761; PIDN:AAA63477.1; PID:q202762  
 R:Perez, D.M.; Piasick, M.T.; Graham, R.M.  
 Mol. Pharmacol. 40, 876-883, 1991  
 A>Title: Solution-phase library screening for the identification of rare clones: isol  
 A:Reference number: A53280; MUID:92100054  
 A:Accession: A53280  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-36, 'P', 38-58, 'I', 60-202, 'I', 204-305, 'R', 307-366, 'I', 368-370, 'I', 372-559  
 A:Experimental source: hippocampus  
 A>Note: sequence extracted from NCBI backbone (NCBIP:73541)  
 C:Superfamily: vertebrate rhodopsin  
 C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein

Query Match	Score	DB 2;	Length	560
11.68;	306.5;			

Best Local Similarity 23.08; Pred. No. 3.1e-13;

Matches 102; Conservative 59; Mismatches 163; Indels 119; Gaps 10;

29 AHGIRSTVLVIFLASFGNIVLALVQRKPÖLLQVTNRFIENLLVTBLLQISLVAPW 88

D 88 A Q G V G V G F L A A F I L I A V A G N L L V I L S V A C N R H L Q I V I N F I V N L A V A D L L L S A A V L P E S 14

89 VAISVPLFWPLNSHF:CIALVSLIHLFAASVNI:VLVSVDRI:LSL:HP:LS:PS:NM:UR:RG: 14

148 AIMEVLOFWAFGRICDVMARVDLCLIASLSLPLISVDNIVGVNBSLIFALMIENNA 20

[illegible][illegible][illegible]

—  
..

—  
..

..

..

—  
—  
—

—  
—

—

[illegible]

310	=====	323
	WATER FACT CURT	FRECEDE
	:    :	:     :

300 FI BBBCVOCRAAKVTTTTTTCVVVCI CBBVCBI BVI AVTUDVETOVBOVTTTTIWI EEI OC 44

[illegible]







DR PROSITE; P500237; G\_PROTEIN\_RECPT\_FL1; 1.  
DR PROSITE; P550262; G\_PROTEIN\_RECPT\_FL2; 1.  
DR G-protein coupled receptor; Receptor; Transmembrane  
SEQUENCE 455 AA; 50802 MW; 80C60918FA08B6C CRC64;

Query Match	12.8%	Score 338.5;	DB 4;	Length 455;
Best Local Similarity	21.3%;	Pred. No. 8.9e-18;		
Matches 108; Conservative	80;	Mismatches 187;	Indels 131;	Gaps 11;

QY	8	STRESNSHNCMPKSKPISIAHIIISVTLVFLASFGNIVLAVLVDORRPOLOWYN	67
Db	5	SGNMSDSNCTOP--PAPVMSKAIIILGVLIGLIEFGVGNITLVILSVACHHLLSYTH	62
QY	68	RFIFNLVTLDLQISLVAPMVAVTSVLEPMLNSHECTALVSLTHLFAFASVNTVLVSV	127
Db	63	YYIYNLAVADLLLSVLPSPSAIFEVLYGWAFCRGVFCNIMAAVDVLTCTASINGLIIISL	122
QY	128	DRVSIITHPLSPYSKMTORGQYLLLYGTWTLVAILQSTPPLYGWGQAIFDRMLAGSIMWG	187
Db	123	DRIVGVSFPLRPIYIVQRGMLALICVMAISLVISIGLPLFGMRQRA--PEDDETTCOI--N	179
Y	188	ASPSYTLISVVSFVLPIELIWMACYSVFCAARRQHALLYNKRHSLEVRVKDCVENEDE	247
Db	180	EEEPYVFEFSAIGSEYTLPLAIIILWYCVYVYAKR-----ESRGL	218
QY	248	EGAKKKEEFODESEFRQHGHEVYAKKGRMEANDGSLAKESGTCSTSESSVEARGSEVR	307
Db	219	KSGIKTKSDSEOVYTLIHRKNAPA-----GSGMMSATKTHFSVRL--	261
QY	308	ESTVASDGSMEGKEGSTKVEENSMKADKGRTEVNOCSIDLGEDDMFEGEDDINFSEDDV	367
Db	262	-----LKFNRK--	268
QY	368	EAVNIPESLPPSRKNSNSNPPLPRCYOCKAKAYIFIIITSEYVLSLGPCLAVLAVAVDY	427
Db	269	-----KAKTLGIIVGCFVLCMPF--FLVMPISFFP	299
QY	428	ETVQPVQWITIIILFLQCCIHPRYYGVYMKHTIKREIOMLK--KFPCKEKPKPE-----D	482
Db	300	DFKRSSEIVFKYFVWLGLNSCINPIIYPCSSQERKKAFOVNLRIQCLCRQSSKHALGYT	359
QY	483	SHPLPCTEGTEGTEKI--VPSTYDATF	507
Db	360	LHPPSAQVEGQHKDMVRIIPGSRRETF	385

```

RESULT 11
Q9N296                                PRELIMINARY;      PRT;      422 AA.
ID   Q9N296
AC   Q9N296;
DT   01-OCT-2000 (TREMBLrel. 15, Created)
DT   01-OCT-2000 (TREMBLrel. 15, last sequence update)
DT   01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE   SEROTONIN RECEPTOR 1A.
GN   HTR1A.
OS   Pongo pygmaeus (Orangutan) .
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Pongo.
CX   NCBI_TaxID=9600;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OR-POL3;
RA   Kitano T., Kobayakawa H., Saitou N.;
RT   "Silver Project";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY
CC   -|- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTOR
DR   EMBL; AB041406; BAA94491.1; -.
DR   HSSP; P29274; 1MMH.
DR   InterPro: IPR000276; GPCR_Rhodpsn.
DR   Pfam: PF00001; 7tm1.1.
DR   PRINTS; PR00237; GPCRHHODPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECPT_FL1; 1.

```

DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane  
SQ SEQUENCE 422 AA; 46121 MM; C2CCC9803E8F8F9B CRC64;

Query Match	12.7%	Score 335.5;	DB 6;	Length 422;
Best Local Similarity	22.5%;	Pred. No. 1.4e-17;		
Matches 104; Conservative	89;	Mismatches 187;	Indels 83;	Gaps 10;

```

Oy 26 ISAHGIRSRVNLVFLAASVNG--LYLVLRORQQLQVNRFFENLVTDLQISL 83
Db 31 VTFESQVYVTLSTLGLTFPCALVACVAAVAAIALER--SLQNVANYLIGSLAVDLMVSVL 88
Oy 84 VAPWVATSVBLEWPLNSHFCTALVSLTHLEFAASVNTVLVSVDRYLSIINPLSPSKM 143
Db 89 VLPMAALVQVINKMTLGGVTCDFLEFALDVLCSTCSLHLICALDRLVMAITDPRIDYVNRK 148
Oy 144 TORQYLLLYGTWIVAILQSTPRPLXGMOAARFDBRNALCSIMGASSTYTLISVSTVI 203
Db 149 TPRRAALISLTWLIGFLISLIPMLGWTP--EERSDDPDCSTISKDHGYTYVTFGAFTI 206
Oy 204 PLIVAIACYSVFCARQNALHTLVNKRHSLEVBVKCQVENEDEGA-----EKK 253
Db 207 PLMLLVLYGSIIFRAR-----PRIKTYVKKVEGTAGDTHNGASPARQK 251
Oy 254 EEFODESEFRQHEGEVYAKAGTGRMEADGSLAKAEGSTGTSESSVEARGSEBEVRESSTVA 313
Db 252 KSVNSESSESRMMRLG-VESKAGGGLCANGAV--RQGDGALAEVIEVH-----296
Oy 314 SDGSMESGEGSTKYEENSMKADKRTFENQCSIDLGEDMFEFGDDIINFSEDOEVANIP 373
Db 297 -----RVGNSK-EHLPLPSEAGPTPCAPASFE-----322
Oy 374 ESLPFSRNSNSNPEPLPCYOCKAKAYLFIIFSVYSLGVCFLAVLAVWVDVETQVQ 433
Db 323 ---RKNEENAEKRRKMALAREKTVKTLGILMGFIILCOMLPEFVALVLPCESSCHMPT 379
Oy 434 WYTIITLFLQCCIHPEYVYGYMHKTKITKEDIDMLKKFPCKE 476
Db 380 LIGALINMLGYSNSLNLNVIAVYENKQOFQNAFKKILCKCFERQ 422

```

```

RESULT 12
ID 013729
AC 013729; PRELIMINARY; PRT; 429 AA.
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE ALPHA 1C ADRENERGIC RECEPTOR ISOFORM 3.
OS Homo sapiens (Human) .
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PROSTATE;
RA Tanaka T.;
RL Submitted (JUL-1994) to the EMBL/Genbank/DBDj databases.
RN [2]
RP SEQUENCE OF 424-429 FROM N.A.
RC TISSUE=PROSTATE;
RA MEDLINE=95255557; PubMed=7737411;
RA Hirasawa A., Shibata K., Horie K., Tanaka T.,
RA Muramoto N., Takagaki K., Yano J., Tsujimoto G.;
RT "Cloning, functional expression and tissue distribution of human alpha
RT 1C-adrenoceptor splice variants."
RL FEBS Lett. 363:256-260(1995).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY) .
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; D34201; BA006900.1; -;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHOOPSPN.
DR

```





Db	310	TWLSFSAVACHPLIIGYIMNKIVTKRELLGM	338
Db	310	TWLSFSAVACHPLIIGYIMNKIVTKRELLGM	338
RESULT 4			
Q96RE8	PRELIMINARY:	PRT:	466 AA.
AC	Q96RE8:		
DT	01-DEC-2001 (Tremblrel. 19, Created)		
DT	01-DEC-2001 (Tremblrel. 19, Last sequence update)		
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)		
DE	ADRENERGIC RECEPTOR ALPHA-1A.		
GN	ADRA1A.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_taxid:9606;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RA	Banerjee A.G.N., Aarti A.		
RT	"R1-PCR cloning and sequence analysis of adrenergic receptor subtype-		
RT	alpha-1a cDNA from human prostatic cell-line DU-145."		
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: AF395806; AKK7197.1; ..		
KW	Receptor.		
SO	SEQUENCE	466 AA; 51431 MW; 876CBB3E32B7A1 CRC64;	
Query Match	13.0%; Score 343; DB 4; Length 466;		
Best Local Similarity	21.6%; Pred. No. 4,1e-18;		
Matches	106; Conservative 76; Mismatches 178; Indels 130; Gaps 10;		
QY	8	STRESNSSHTCMPLSKMPISLAHGIIIRSTVLYFLAASVGVNIVLALYQRRPOLQYTN	67
Db	5	SENASTSSNCTOP--PAPVNISSKAILGLVILGLLFGVLGNILVLSVACHRHLSVTH	62
QY	68	REIFMLVTLDIQISLVAVWVATSVLPVPLNSHCTALVSLTHFAFASVNTIVLV	127
Db	63	YIVVLAVADELITSTVLPFSALFEYLVGMAGRGFCNIMAAVDVLCCTASIMGLCTISI	122
QY	128	DYLSIIHPLSYPSKMTORRGYLLYTWIVAILLOSTPPLVGMGQAAPDERNALCSMIMG	187
Db	123	DYIGVSHPLRPTLYTQRGMLALCVWALSLVLSIGLPFGMRQPA-PEDETICQI--N	179
QY	188	ASPSYTLISVSFYIPLVIMTACYSVVFCAARQHALLYNVKRSLEVRKDCVENEDE	247
Db	180	EEPGVYLFALSGFLPLAILIIMYCRVYVAKR-----ESRGL	218
QY	248	EEAEKKEEQDSEERRQHEGEVAKKEGMEAKDGSIRAKESSTGTSESSVEARGSEVR	307
Db	219	KSGLTIDKSDSQVTLRIHRKNAP-----GSGASAKTYTHRSVRL-	261
QY	308	ESSYVAADSGMEKGEGSTVKEENSMKADKPTENVNOCSIDLGEDDMEFEDDINSESDV	367
Db	262	-----LKSREK-	268
QY	368	EAVNIPESLLPSRRSNSNPPLRCYQCKAAKAVFIILIFSYYLSLGPYCEFLAVLAVWDV	427
Db	269	-----KAAKTLIGVGCFLWMLPF-FLVMPISGFPP	299
QY	428	ETVOGQWVTTIIMLFLOCCIHPIVYGVGMHTIKKEIDMKL-KFCKEKRPKE---D	482
Db	300	DKPSETVYKIVFWLGYLNSCINPIIYPCSSQSEFFKAFQNVLRIOCLCKRQSSKHALGYT	359
QY	483	SHPDLPTEG	492
Db	360	LHPSPQAVEG	369
RESULT 5			
Q9N298	PRELIMINARY:	PRT:	422 AA.
AC	Q9N298:		
DT	01-OCT-2000 (Tremblrel. 15, Created)		

DT	01-DEC-2000 (TREMBLrel. 15, last sequence update)
DT	01-OCT-2000 (TREMBLrel. 19, last annotation update)
DE	SEROTONIN RECEPTOR 1A.
GN	HTR1A.
OS	Pan troglodytes (Chimpanzee).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Pan.
OX	[NCBI_TaxID=9598;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-CHIMP-220;
RA	Kitano T., Kobayakawa H., Saitou N.;
RT	"Silver Project";
RL	Submitted (Apr-2000) to the EMBL/Genbank/DDBJ databases.
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR	EMBL; AB041404; BA049489.1; -
DR	HSSP; P29274; 1MMH.
DR	InterPro: IPR000276; GPCR_Rhodopsn.
DR	Pfam: PF00001; 7tm_1; 1
DR	PRINTS; PR00237; GPCRRHODOPSN.
DR	PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR	PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW	G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SO	SEQUENCE 422 AA; 4611 MW; FCD989BD0313A1A0 CRC64;

  

Query Match	12.9%; Score 340.5; DB 6; Length 422;
Best Local Similarity	22.0%; Pred. No. 5,7e-18;
Matches 101; Conservative 91; Mismatches 191; Indels 77; Gaps	

  

QY	26	ISLAGIIRSYLVVIFFLASPGVN--IVATLVQRKPOLQVTNRITFNLMTDILLQLSL	83
Dd	31	VVFYSQVITISLLGLTFLFCFAVLGMCVVAIALER--SLQANANYLTIGSLAVTDLMWSYL	88
QY	84	VAPWVAATSVPLFPFPLNSHFCTALVSILTFAFASYNTIVLSVDRYLSIIHPLSYPSKM	143
Dd	89	VLPMAALGVCLKWTILGYTTCDFIALDVLCCTSSLHLCAITALDYKAITDPIDVYNKR	148
QY	144	TGRRGYLLTGTVIAVALLOSTPPRLXWGQAADFERNALCSMTIGASPSTTILSVSEFIY	203
Dd	149	TPRRAAALLSTMLIGFLTISPMLGMWRP--EDRSDDPACTISKDHGTIVSTFGAFYI	206
QY	204	PLTWIACISVVCARRQHALLYNWKRLSELYRVKDCEYENDEBEAAEKKEPFQDSSEFR	263
Dd	207	PLLMLIVTLGRFFRAR-----FRIRKYTKVEKTADADRHGASPAQOKP	251
QY	264	ROHEGEVKAKEEIRM--EAKDGSL----KAKEGSTGSSEYEARGSEVRRESSTVASDG	316
Dd	252	KSYNGESGSRMRKLGVESAAGCALCANGAVROGDGALEIVFH-----	296
QY	317	SMEKGEGSTKVEENSNAKADGRTEVNQCSDIDGEDDMEEGEDDINFSEDVEAVNIPESL	376
Dd	297	---RVGNKS-K-EHLPLPSAGPTPCAAPSFE-----	322
QY	377	PPSRNSNSNPPLPRCYQCKAAKVFTITTFYSVLSGPYCFELAVLAWVDVEIQVQWVI	436
Dd	323	RKNERRAEKKRMALARREKRTVKTGIINGTFTILCLMPFIYALVLPFCSSSCHMPTLLG	382
QY	437	TIIMLFELQCIHPHYVGVMKTKIKELLQDIOMLKRFCKE	476
Dd	383	AIIMNLGYSNSLIIPVITAYFMKDFONAEKFKTIKCFCFO	422

  

RESULT 6	
O9MZU2	
ID	O9MZU2 PRELIMINARY; PRT: 425 AA.
AC	O9MZU2:
DT	01-OCT-2000 (TREMBLrel. 15, Created)
DT	01-OCT-2000 (TREMBLrel. 15, last sequence update)
DT	01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE	ALPHA 1A-ADRENORECEPTOR ISOFORM 3.
OC	Oryctolagus cuniculus (Rabbit).
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Db 121 TIVVSVDRYLIIHPLSYSKMTQRRGYLLLYGTWIVAILQSTPPLVYGQAADERNA 180
QY 181 LCMWINGASPTIILSVSTIVPIIYMIACYSVFCARROHALLVNRHSLVEVKD 240
   |||||
Db 181 LCMWINGASPTIILSVSTIVPIIYMIACYSVFCARROHALLVNRHSLVEVKD 240
QY 241 CVENEDDEGAKEKEFEODESEFRROHGEVAKEGMEAKDGLAKEGSTGTSSEVVA 300
   |||||
Db 241 CVENEDDEGAKEKEFEODESEFRROHGEVAKEGMEAKDGLAKEGSTGTSSEVVA 300
QY 301 RGESEFRESSTVSDGSMGEKSTKVEENSMKADKRTVENOCSDLGEDDMFEGDDI 360
   |||||
Db 301 RGESEFRESSTVSDGSMGEKSTKVEENSMKADKRTVENOCSDLGEDDMFEGDDI 360
QY 361 NFSEDDVEAVNIPESLPPSRNSNPRLPCRYOCKAKVFIIFISYVLSIGYCFCLAV 420
   |||||
Db 361 NFSEDDVEAVNIPESLPPSRNSNPRLPCRYOCKAKVFIIFISYVLSIGYCFCLAV 420
QY 421 LAWVVDVETQVPMVITIIIMLFLOCCIHPIYVYGVMHKTIKKEIDMLKFFCKEPRK 480
   |||||
Db 421 LAWVVDVETQVPMVITIIIMLFLOCCIHPIYVYGVMHKTIKKEIDMLKFFCKEPRK 480
QY 481 EDSHPDLPTGEGTEGKIYPSYDSATRP 508
   |||||
Db 481 EDSHPDLPTGEGTEGKIYPSYDSATRP 508

```

## RESULT 2

```

Q90X46 PRELIMINARY; PRT; 402 AA.
ID Q90X46
AC Q90X46;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE SC:B2015.4 (NOVEL PROTEIN SIMILAR TO HUMAN G-PROTEIN COUPLED RECEPTOR
DE R22).
GN SC:B2015.4.
OS Brachydanio rerio (zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_taxid=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd D.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL590146; CAC94897.1; -.
KW Receptor.
SQ
SEQUENCE 402 AA; 44373 MW; 5FB8875874267F00 CRC64;

```

```

Query Match 14.88; Score 392.5; DB 13; Length 402;
Best Local Similarity 23.88; Pred. No. 5.9e-22;
Matches 105; Conservative 72; Mismatches 127; Indels 138; Gaps 9;

```

```

QY 32 IIRSTVLVIFLASFGVNIYALVLRKPOLQVYTNRFIFNLVTDLOISLVAIPVAV 91
   |||||
Db 24 VLESVSIITAIACLGNIIVVTLVKKPYLTPSNKFVSLTSLNLLSLVAMPVVAS 83
   |||||
QY 92 SVLPFWPLNSHFC--TALVSLTHLFAFASVNTIVLVSDRLSTIHPLSYSKMTQRRG 149
   |||||
Db 84 SVRRDMFVGMVCMFTALHL--LVSSSMULTGAIADRYAVLPIYPMKITGNRAV 141
   |||||
QY 150 LLATGTIVAILOSTPPLVYGQAADERNALCSMIGASPTIILSVSTIVPIIYMI 209
   |||||
Db 142 LAIYIMLHSLVGLPLPLFGSSPEFDRFKWTCTVSMHKEISYAFVWTWCCLPLVAM 201
   |||||
QY 210 ACSVVFCAARROHALLVNRHSLVAVKDCVENEDEGAKEKEFEODESEFRROHGE 269
   |||||
Db 202 VCYGVIFRRVAR----- 212
QY 270 VKAEGMEAKDGLAKEGSTGTSSEVFAKSGSEVRESSTVSDGSMGEKSTKVEE 329
   |||||
Db 213 -----IKARKVYCG-----SVVVS-----QE 228

```

```

QY 330 NSMKADKRTVENOCSDLGEDDMFEGDDINFSEDDVEAVNIPESLPPSRN---SNSN 386
   |||||
Db 229 ESSSQNNGRKNSNTSTSSG-----SRKSLITSGS- 258
   |||||
QY 387 PPLPRCYOCKAKAVFIIFISYVLSIGYCFCLAVLAWVVDVETQVPMVITIIIMLFLO 446
   |||||
Db 259 -----QCKAFITIIIVLGLFTWGPVYVISTEALGKNSVSPQ-VETLVSWLSFTS 310
   |||||
QY 447 CCHPIYVYGVMHKTIKKEIDM 468
   |||||
Db 311 AVCHPLIYGLMNTKVRKELLGM 332

```

## RESULT 3

```

ID 075963 PRELIMINARY; PRT; 407 AA.
AC 075963;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE G-PROTEIN COUPLED RECEPTOR R22.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-BRAIN;
RA Raming K.;
RT "Identification of a novel human G-protein coupled receptor.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF091890; AAC61598.1; -.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ
SEQUENCE 407 AA; 45363 MW; 00078BB8B5C1F21 CRC64;

```

```

Query Match 14.88; Score 391. DB 4; Length 407;
Best Local Similarity 23.68; Pred. No. 7.8e-22;
Matches 106; Conservative 67; Mismatches 130; Indels 146; Gaps 11;

```

```

QY 31 GIIRSTVLVIFLASFY--GNIVYALVLRKPOLQVYTNRFIFNLVTDLOISLVAIPW 88
   |||||
Db 25 GVITTOFTAIIVITFEVLGMLVIVTLVKKSVLTLSNKFVSLTSLNLLSVLVLPV 84
   |||||
QY 89 VATSVPLFWPLNSHFCYALVSLTHLFAFASVNTIVLVSDRYLSIHPLSYSKMTQRRG 148
   |||||
Db 85 VTSSIRREMFVGMVCMFNSALYLLISSASMLTIGVLAIDRYAVLPLVPMYPMKITGNRA 144
   |||||
QY 149 YLLIYGTWIVAILOSTPPLVYGQAADERNALCSMIGASPTIILSVSTIVPIIYMI 208
   |||||
Db 145 VMAIVYIMLHSLICLPLPLFGSSPEFDRFKWCMVAAHRRPGYTAQWQICALFPLVM 204
   |||||
QY 209 IACSVVFCAARROHALLVNRHSLVAVKDCVENEDEGAKEKEFEODESEFRROHGE 268
   |||||
Db 205 LVCTGIFRRVAR-----VKARKVHCGYIYVE-EDAQ----- 235
   |||||
QY 269 EVKAKEGMEAKDGLAKEGSTGTSSEVFAKSGSEVRESSTVSDGSMGEKSTKVE 328
   |||||
Db 236 ----RTGR-----KNSSTSTSSG----- 250
   |||||
QY 329 ENSMKADKRTVENOCSDLGEDDMFEGDDINFSEDDVEAVNIPESLPPSRNS 383
   |||||
Db 251 -----SRAAFQGV 260
   |||||
QY 384 -NSNPPLPCRYOCKAKAVFIIFISYVLSIGY-CFLAVLAW--VDVETQVPMVITII 439
   |||||
Db 261 YSAN-----QCKALTIILVLCAPVWTGPIVAVLASEALGKSSVPSLETWA----- 309
   |||||
QY 440 IMLFLOCCIHPIYVYGVMHKTIKKEIDM 468

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 16:42:43 (Search time 31 Seconds  
(without alignments)  
2834.885 Million cell updates/sec

Title: US-09-992-238-2

Perfect score: 2644  
Sequence: 1 MTSCTNSTRESNSHTCMP.....GTEGTEGKIIVPSYDSATFP 508

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: SP\_TREMBL\_19:\*  
2: SP\_archaea:\*  
3: SP\_bacteria:\*  
4: SP\_fungi:\*  
5: SP\_human:\*  
6: SP\_invertebrate:\*  
7: SP\_mhc:\*  
8: SP\_organelle:\*  
9: SP\_phage:\*  
10: SP\_plant:\*  
11: SP\_rodent:\*  
12: SP\_virus:\*  
13: SP\_vertebrate:\*  
14: SP\_unclassified:\*  
15: SP\_rvirus:\*  
16: SP\_bacteriap:\*  
17: SP\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2641	99.9	508	4	Q96P66
2	392.5	14.8	402	13	Q90X46
3	391	14.8	407	4	Q75963
4	343	13.0	466	4	Q96R68
5	340.5	12.9	422	6	Q9N298
6	340	12.9	425	6	Q9N203
7	340	12.9	429	6	Q9N202
8	339.5	12.8	408	13	Q98998
9	339.5	12.8	422	6	Q9N297
10	338.5	12.7	455	4	Q60451
11	335.5	12.6	422	6	Q9N296
12	333.5	12.6	429	4	Q13729
13	333.5	12.6	499	4	Q13675
14	325.5	12.3	466	4	Q9UD63
15	325.5	12.3	466	4	Q9T5W7
16	324	12.3	466	6	Q9T5W7

17	317	12.0	571	6	Q9T5W9	Q9T5W9 sus scrofa
18	304.5	11.5	601	5	Q95YF4	Q95YF4 drosophila
19	300.5	11.4	559	11	Q9GW71	Q9GW71 rattus sp.
20	299	11.3	631	5	Q9VEG1	Q9VEG1 drosophila
21	297.5	11.3	349	4	Q9WEG1	Q9WEG1 homo sapien
22	291.5	11.0	461	13	Q9WMO7	Q9WMO7 oncorhynch
23	291	11.0	474	4	Q96R68	Q96R68 homo sapien
24	287	10.9	456	4	Q96R68	Q96R68 homo sapien
25	285	10.8	610	5	Q44986	Q44986 caenorhabdi
26	283	10.7	515	11	Q9DBL0	Q9DBL0 mus musculi
27	282.5	10.7	432	13	Q9OWY6	Q9OWY6 brachydanio
28	280	10.6	391	5	Q96716	Q96716 brachydanio
29	280	10.6	510	13	Q9OWY5	Q9OWY5 brachydanio
30	279.5	10.6	419	5	Q77254	Q77254 boophilus m
31	279	10.6	518	6	Q9MYR8	Q9MYR8 oryctolagus
32	276	10.4	440	6	Q9N2A5	Q9N2A5 pongo pygma
33	276	10.4	445	4	Q9N2R3	Q9N2R3 homo sapien
34	275.5	10.4	526	4	Q96RG7	Q96RG7 homo sapien
35	275	10.4	528	13	Q9PTF6	Q9PTF6 homo sapien
36	274	10.4	370	4	Q9H1N4	Q9H1N4 gallus gall
37	274	10.4	440	4	Q9P1X9	Q9P1X9 homo sapien
38	274	10.4	440	6	Q9N2A7	Q9N2A7 pan troglod
39	274	10.4	440	6	Q9N2A6	Q9N2A6 gorilla gor
40	274	10.4	443	4	Q9UPA9	Q9UPA9 homo sapien
41	272	10.3	477	5	Q9BMA9	Q9BMA9 mamestra br
42	270	10.2	470	11	P97842	P97842 rattus norv
43	267.5	10.1	396	4	Q9NOS5	Q9NOS5 homo sapien
44	266	10.1	392	13	Q9YHA5	Q9YHA5 myxine glut
45	265.5	10.0	414	6	Q9GK99	Q9GK99 canis famli

## ALIGNMENTS

RESULT 1  
ID Q96P66 PRELIMINARY; PRT; 508 AA.

AC Q96P66; 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE G PROTEIN-COUPLED RECEPTOR.  
GN GPR101.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21458557; PubMed=11574155;  
RA Lee D.K., Nguyen T., Lynch K.R., Cheng R., Vantl W.B., Arkhicko O.,  
Lewis T., Evans J.F., George S.R., O'Dowd B.F.;  
RT "Discovery and mapping of ten novel G protein-coupled receptor  
genes".  
RL Gene 273:83-91(2001).  
DR EMBL; AF411115; AAL26486.1; .  
KW RECEPTOR.  
SQ SEQUENCE 508 AA; 56716 MW; E20A409P65C95B85 CRC64;

Query Match 99.9%; Score 2641; DB 4; Length 508;  
Best Local Similarity 99.8%; Pred. No. 4.9e-193;  
Matches 507; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTSCTNSTRESNSHTCMPISLAHGIIRSTVLVIFLAASFVGNVLAIVLRKP 60  
DB 1 MTSCTNSTRESNSHTCMPISLAHGIIRSTVLVIFLAASFVGNVLAIVLRKP 60  
QY QLOVYTRFFINLVTDLOISYAPWVATSVLPFLNHFCAVSLTHLFAFASVN 120  
DB QLOVYTRFFINLVTDLOISYAPWVATSVLPFLNHFCAVSLTHLFAFASVN 120  
QY QLOVYTRFFINLVTDLOISYAPWVATSVLPFLNHFCAVSLTHLFAFASVN 120  
DB QLOVYTRFFINLVTDLOISYAPWVATSVLPFLNHFCAVSLTHLFAFASVN 120  
QY TIVAVSVDRYLSIHPPLSPKMTQVRGTYLLGTWIVAIQSPPLVGVGQAAFDENNA 180  
DB TIVAVSVDRYLSIHPPLSPKMTQVRGTYLLGTWIVAIQSPPLVGVGQAAFDENNA 180





```

QY 180 AICSMWASPSYTLISVSVFTIPLIWIACYSVFCARROHALLVVKRHSLEVRK 239
DB 244 GTC--LISODKGYTIFSTVGAFYLPMLVMIITIRLWLABSR-----IRDKQOMKRA 295
QY 240 DCEVNEDEGAKKEEFODESEFRQHEGYAKAKGRMEA-----KDSLIK 285
DB 236 RLKTEETLVASPKTEYSVSDCNCNSPDSTTEKKRRAPKSYGSCPREEPRKKNRAK 355
QY 286 AKEGSTGSESSVEARGSEEV-ESSIVASDGSMEGSGSTVEENSKADGREFVNOCC 344
DB 356 LPERANGVNSNSSSERLKQIETAEAFANGCAE--EASTAMLE-----RQC 401
QY 345 SIDGEDMEFGEEDDINFSEDDVEAVNIPESLPPSRNNSNPNPLRCYOCKAKAVIFLI 404
DB 402 N-----NGKRTISSNDTPYSRTREKLEKRR-----KAARTLAI 436
QY 405 IFSVYLSGPGYCLAVLAWVDVEYQVQWITITITIFLQCCIHPIVYGMHTIKE 464
DB 437 TGAFLICWLPFFITALISGPDPE-GIPFARSFVLMWGYNSLNPITITIFSPERSA 495
QY 465 IQDML 469
DB 496 FOKIL 500

RESULT 15
ALAD_MOUSE STANDARD: PRT: 562 AA.
ID ALAD_MOUSE
AC P97714; 061619;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-1D adrenergic receptor (Alpha 1D-adrenoceptor) (Alpha-1A
adrenoceptor).
GN ADRAD1 OR ADRAD1A OR GPCR8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=56064818; PubMed=7595531;
RA Alonso-Llamazares A., Zamarrillo D., Casanova E., Ovalle S., Calvo P.,
Alucheiru M.A.;
RT "Molecular cloning of alpha 1d-adrenergic receptor and tissue
distribution of three alpha 1-adrenergic receptor subtypes in
mouse."
RT J. Neurochem. 65:2387-2392(1995).
RN [2]
RP SEQUENCE OF 190-350 FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=54116980; PubMed=8288218;
RA Wilkie T.M., Chen Y., Gilbert D.J., Moore K.J., Yu L., Simon M.I.,
Copeland N.G., Jenkins N.A.;
RT "Identification, chromosomal location, and genome organization of
mammalian G-protein-coupled receptors."
RT Genomics 18:175-184(1993).
CC -1- FUNCTION: THIS ALPHA-ADRENERGIC RECEPTOR MEDIATES ITS EFFECT
THROUGH THE INFUX OF EXTRACELLULAR CALCIUM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S80044; AAB47042.1; -

```

```

DR EMBL; L20333; AAB16854.1; -
DR GCRDB; GCR 0810; -
DR GCRDB; GCR 1976; -
DR MGD; MGI:106673; Adra1d.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHDOPSN.
DR PROSITE; PS00237; G-PROTEIN_RECEP_FL_1; 1.
DR PROSITE; PS0262; G-PROTEIN_RECEP_FL_2; 1.
KW G-protein coupled receptor; transmembrane; glycoprotein;
KW Multigene family; phosphorylation; lipoprotein; palmitate.
KW Domain
FT DOMAIN 1 90
FT TRANSMEM 91 115
FT DOMAIN 116 127
FT TRANSMEM 128 153
FT DOMAIN 154 163
FT TRANSMEM 164 186
FT TRANSMEM 187 207
FT DOMAIN 208 232
FT TRANSMEM 233 245
FT TRANSMEM 246 269
FT DOMAIN 270 342
FT TRANSMEM 343 367
FT DOMAIN 368 374
FT TRANSMEM 375 399
FT DOMAIN 400 562
FT CARBOHYD 60 60
FT CARBOHYD 76 76
FT LIPID 413 413
FT DOMAIN 21 57
FT DOMAIN 416 421
FT CONFLICT 330 330
SQ SEQUENCE 562 AA; 59857 MW; 6CF1515B5F7FA47 CRC64;

```

```

Query Match 12.0%; Score 317.5; DB 1; Length 562;
Best Local Similarity 23.3%; Pred. No. 2e-11;
Matches 103; Conservative 60; Mismatches 161; Indels 119; Gaps 10;

```

```

QY 29 AHGIRSVLYVFLAASVGNIVLALVQKRPOLLQVNRRTFNLVTDLIQISLVAPWV 88
DB 88 AAGVGVGVLAFILTAAGNMLVILSVACNRHLOTVNYFIVMLAVALLSAVALPFS 147
QY 89 VATSVPLFPLNSHFCALVSLTHLFAFASVNTIVLSVDYLSIHPISPSKMTORBG 148
DB 148 ATMEVLGWPGRGRTCDVMAVDVLCASLISCTISVDYVGRHSKTPAINTERKA 207
QY 149 YLLYGTVIVALIILSTPPLYGWQ-AAFDERNALCSMTIGASPSYTLISVSVFTIPLIV 207
DB 208 AAILLMAVALVAVSGVLLGKMKKEVPEDER--FCGIT--EYGVAFPSVCSFYLPMAV 263
QY 208 MIACYSVFCARROHALLVVKRHSLEVRKDCVENDEGAKKEEFODESEFRQHE 267
DB 264 IVVMICRYVAVARS-----TRTSLGATK-----REP 290
QY 268 GEVAKKEGMEAKDGLKAKEGSTGTSSEVARGSEEVRESSIVASDGSMEGSGSTKY 327
DB 291 GKASEVILRIHRCARATSAK-GNGETQSS----- 318
QY 328 EENSMAKDKRGTEVNOCSIDLGEDMEFGEEDDINFSEDDVEAVNIPESLPPSRNNSNP 387
DB 319 -----KGTILSLSSVRL-----LAFSRRK----- 338
QY 388 PLPRCYOCKAKAVIFITIFSVLSLGPYCLAVLAWVDVEYQVQWITITITIFLQCC 447
DB 339 -----KAAKTLAIVAGVFLCWFPEFFVPLGS-LFPOLKPSGEVFKVIFWLYGFS 389
QY 448 CIHPYVYGMHTIKKEIQDMLK 470
DB 390 CVNPLIIPCSSEKRRKRAFLRLR 412

```

Search completed: October 28, 2002, 16:45:59  
Job time : 17 secs

DR PROSITE; PS00237; G-PROTEIN-RECEP\_F1\_1; 1.  
 DR PROSITE; PS50262; G-PROTEIN-RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 Multi-gene family.  
 FT DOMAIN 1 33  
 FT TRANSSEM 34 59  
 FT DOMAIN 60 70  
 FT TRANSSEM 71 95  
 FT DOMAIN 96 106  
 FT TRANSSEM 107 129  
 FT DOMAIN 130 149  
 FT TRANSSEM 150 176  
 FT DOMAIN 181 193  
 FT TRANSSEM 194 219  
 FT DOMAIN 220 339  
 FT TRANSSEM 340 361  
 FT DOMAIN 362 372  
 FT TRANSSEM 373 397  
 FT DOMAIN 398 416  
 FT DISULFID 106 189  
 FT CARBOHYD 5 6  
 FT CARBOHYD 6 6  
 FT CARBOHYD 18 18  
 FT SEQUENCE 416 AA; 47031 MW; 9B4CC413BEC750FE CRC64;  
 Query Match 12.1%; Score 320; DB 1; Length 416;  
 Best Local Similarity 23.0%; Pred. No. 1e-11;  
 Matches 112; Conservative 83; Mismatches 200; Indels 92; Gaps 14;

OY 6 TNSR-----ESNSHTCMPLSKMPSIAHGIRSVIVIFLAASVGN--ITVALYLQR 58  
 DB 4 TNNVTGHTFSTSNRTSKSFE-EVKLSYQVTSFLGALILCSIGMACVVAIALER 62  
 OY 59 KPOLQVTRNRFNLTDLQISLVAPWVAVTSVPLFPLNSHFCALVSLTHLFAS 118  
 DB 63 --SLQNVANYLIGSLAVTDLVSVLPMAALVOYLNRMTGQIPCDIFISLDMCTSS 120  
 OY 119 VNTVLVSVDRYLSIIHPLSPKMTQRGYLLGTWIVAILLOSTPLVGM---GOAA 174  
 DB 121 ILHCVIALDRYMAITEPIDMKRTPRRAVLISVTMLVFESISIPMLMRQSPSSMA 180  
 OY 175 FDERNA-LCSMTWIGASPTITLVSFIVPLIWMACYSVFCAARRQHALLYVKRHS 233  
 DB 181 EDRANSKCKCKIT--QDWYTIYSTFGAFYIPLTMLVLYGRIFKAAR----- 225  
 OY 234 LEVAVKOCVENDEGAKEKEPDESEFRQHGSEVAKARKREADGSLKAKEGSTGT 293  
 DB 226 --FRIKRTVKRTEKKVSDTCALSPAFPHKTKPGDAHGKSKW----- 266  
 OY 294 SESSVEARGSEEVESSTVASDG--SMEGKEGSTVEENSMKADGRTEVNGCSIDLGED 352  
 DB 267 --RVEPRPLPVNGAVKHAHGESLD---IIEVQNS-----RCNPL----- 305  
 OY 353 MEFEEDINFSDDVEAVNPESLP---PSRRNSNSNPLPCYQCKAAKAVIFITIFSY 408  
 DB 306 -----PNEPGTVLPFENRHEKATEFETKRIALAREKRTVKTALIMTGF 348  
 OY 409 VLSGPCFLAVLAVWVDEVQVPOWVTTITIMLFLOCCIHPRVYVCGMHTTIKEIDM 468  
 DB 349 ILCLLPFIYALVLPFOESCCEPMHMLKDVIMLWGLSYNSLLPITAYAVENKDFOSAKRK 408  
 OY 469 LKRFCK 475  
 DB 409 IKCHFCK 415

RESULT 14  
 5HT\_LYMT  
 ID 5HT\_LYMT STANDARD; PRT; 509 AA.  
 AC 025414;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE 5-hydroxytryptamine receptor (5-HT receptor) (Serotonin receptor).  
 OS Lymnaea stagnalis (Great pond snail).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;  
 OC Lymnaeidae; Lymnaea.  
 OX NCBI\_TaxID=6523;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93126323; PubMed=8093556;  
 RA Sugamori K.S., Sunahara R.K., Guan H.-C., Bulloch A.G., Tensen C.P.,  
 RA Seaman P., Niznik H.B., van Tol H.H.;  
 RT "Serotonin receptor cDNA cloned from Lymnaea stagnalis.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:11-15(1993).  
 CC -1- FUNCTION: THIS IS A RECEPTOR FOR 5-HYDROXYTRYPTAMINE (SEROTONIN),  
 CC A BIOGENIC HORMONE THAT FUNCTION AS A NEUROTRANSMITTER, A HORMONE,  
 CC AND A MITOGEN.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L06803; AAA29290.1; -.  
 DR GCRDB: GCR\_0520; -.  
 DR InterPro: IPR000276; GPCR\_Rhodpsn.  
 DR Pfam: PF00001; 7tm\_1; 1.  
 DR PRINTS: PR00237; GPCR\_Rhodpsn.  
 DR PROSITE: PS00237; G-PROTEIN-RECEP\_F1\_1; 1.  
 DR PROSITE: PS50262; G-PROTEIN-RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 FT DOMAIN 1 99  
 FT TRANSSEM 100 122  
 FT DOMAIN 123 132  
 FT TRANSSEM 133 154  
 FT DOMAIN 155 169  
 FT TRANSSEM 170 191  
 FT DOMAIN 192 210  
 FT TRANSSEM 211 233  
 FT DOMAIN 234 259  
 FT TRANSSEM 260 281  
 FT DOMAIN 282 432  
 FT TRANSSEM 433 456  
 FT DOMAIN 457 465  
 FT TRANSSEM 466 488  
 FT DOMAIN 489 509  
 FT CARBOHYD 3 3  
 FT CARBOHYD 47 47  
 FT CARBOHYD 58 58  
 FT CARBOHYD 68 68  
 FT CARBOHYD 72 72  
 FT CARBOHYD 78 78  
 FT DISULFID 168 246  
 FT SEQUENCE 509 AA; 56902 MW; D283696C8C50B1B8 CRC64;  
 Query Match 12.1%; Score 319; DB 1; Length 509;  
 Best Local Similarity 22.3%; Pred. No. 1.5e-11;  
 Matches 108; Conservative 96; Mismatches 205; Indels 76; Gaps 14;

OY 3 STCNSTRESNTHCMPLSKMPSIAHGIRSVIVIFLAASVGN--ITVALYLQR 60  
 DB 74 TSATNMDDRYWLSLV-----YSHELVLTSTVILGIFVLCITIGNCFYAAVLER-- 124  
 OY 61 QLLQVTRNRFNLTDLQISLVAPWVAVTSVPLFPLNSHFCALVSLTHLFASVN 120  
 DB 125 SLHVVANVLISLAVADLMAVILVPLSVSEISKVWFHSEVCDMMISVDVLCSTASIL 184  
 OY 121 TIVLVSVDRYLSIIHPLSPKMTQRGYLLGTWIVAILLOSTPLVGMQAAD--ERN 179  
 DB 185 HLVAIAMDRYWAVT--SIDYIRRSARILLIMVWIVALEFISIPPLFGWMDPNDDKT 243

[illegible]

```

Db 299 PNEFPETVEKIVFWLGYLNSCINPIIYPCSSQCFKFAFQVNLRIQCLRROSSKHALGY 358
OY 484 --HPDLPGTEGGTEGKI-VPSYDSATF 507
Db 359 TLHPPSOAVEQHRGMVRIPIGSGETF 385

RESULT 11
SH1A_RAT STANDARD; PRT; 422 AA.
AC P19327;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 5-hydroxytryptamine 1A receptor (5-HT-1A) (Serotonin receptor) (5-HT1A).
GN HTR1A OR 5HT1A.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90202832; PubMed=2156831;
RA "Albert P.R., Zhou Q.-Y., van Tol H.H.M., Bunzow J.R., Civelli O.;
RT "Cloning, functional expression, and mRNA tissue distribution of the
RT rat 5-hydroxytryptamine 1A receptor gene.";
RT J. Biol. Chem. 265:5825-5832(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90355775; PubMed=2167416;
RA Fujimura Y., Nelson D.L., Kashiwara K., Varga E., Roeske W.R.,
RT "Role of cytochrome P450 in the control of the production of
RT erythropoietin.";
RT Life Sci. 47:127-132(1990).
CC -1- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR 5-
CC HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC HORMONE THAT FUNCTIONS
CC AS A NEUROTRANSMITTER, A HORMONE, AND A MITOGEN. THE ACTIVITY OF
CC THIS RECEPTOR IS MEDIATED BY G PROTEINS THAT INHIBITS ADENYLATE
CC CYCLASE ACTIVITY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC SPROGESTS TO THE OTHER 5HT-1 SUBTYPE RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J05276; AAA40612.1; -.
DR PIR; A35181; A35181.
DR PIR; JH0315; JH0315.
DR HSSP; P29274; 1MHM.
DR GCRDB; GCR_0109; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPS.
DR PROSITE; PS00237; G-PROTEIN_RECEP_FL_1; 1.
DR PROSITE; PS0262; G-PROTEIN_RECEP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Multigene family.
FT DOMAIN 1 36 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 37 62 1 (POTENTIAL).
FT DOMAIN 63 73 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 74 98 2 (POTENTIAL).
FT DOMAIN 99 109 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 110 132 3 (POTENTIAL).
FT DOMAIN 133 152 CYTOPLASMIC (POTENTIAL).

```

```

FT TRANSMEM 153 178 4 (POTENTIAL).
FT DOMAIN 179 191 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 192 217 5 (POTENTIAL).
FT DOMAIN 218 245 6 (POTENTIAL).
FT TRANSMEM 246 267 7 (POTENTIAL).
FT DOMAIN 268 295 8 (POTENTIAL).
FT TRANSMEM 296 323 9 (POTENTIAL).
FT DOMAIN 324 351 10 (POTENTIAL).
FT TRANSMEM 352 379 11 (POTENTIAL).
FT DOMAIN 380 407 12 (POTENTIAL).
FT TRANSMEM 408 435 13 (POTENTIAL).
FT DOMAIN 436 463 14 (POTENTIAL).
FT TRANSMEM 464 491 15 (POTENTIAL).
FT DOMAIN 492 519 16 (POTENTIAL).
FT TRANSMEM 520 547 17 (POTENTIAL).
FT DOMAIN 548 575 18 (POTENTIAL).
FT TRANSMEM 576 603 19 (POTENTIAL).
FT DOMAIN 604 631 20 (POTENTIAL).
FT TRANSMEM 632 659 21 (POTENTIAL).
FT DOMAIN 660 687 22 (POTENTIAL).
FT TRANSMEM 688 715 23 (POTENTIAL).
FT DOMAIN 716 743 24 (POTENTIAL).
FT TRANSMEM 744 771 25 (POTENTIAL).
FT DOMAIN 772 800 26 (POTENTIAL).
FT TRANSMEM 801 828 27 (POTENTIAL).
FT DOMAIN 829 856 28 (POTENTIAL).
FT TRANSMEM 857 884 29 (POTENTIAL).
FT DOMAIN 885 912 30 (POTENTIAL).
FT TRANSMEM 913 940 31 (POTENTIAL).
FT DOMAIN 941 968 32 (POTENTIAL).
FT TRANSMEM 969 996 33 (POTENTIAL).
FT DOMAIN 997 1024 34 (POTENTIAL).
FT TRANSMEM 1025 1052 35 (POTENTIAL).
FT DOMAIN 1053 1080 36 (POTENTIAL).
FT TRANSMEM 1081 1108 37 (POTENTIAL).
FT DOMAIN 1109 1136 38 (POTENTIAL).
FT TRANSMEM 1137 1164 39 (POTENTIAL).
FT DOMAIN 1165 1192 40 (POTENTIAL).
FT TRANSMEM 1193 1220 41 (POTENTIAL).
FT DOMAIN 1221 1248 42 (POTENTIAL).
FT TRANSMEM 1249 1276 43 (POTENTIAL).
FT DOMAIN 1277 1304 44 (POTENTIAL).
FT TRANSMEM 1305 1332 45 (POTENTIAL).
FT DOMAIN 1333 1360 46 (POTENTIAL).
FT TRANSMEM 1361 1388 47 (POTENTIAL).
FT DOMAIN 1389 1416 48 (POTENTIAL).
FT TRANSMEM 1417 1444 49 (POTENTIAL).
FT DOMAIN 1445 1472 50 (POTENTIAL).
FT TRANSMEM 1473 1500 51 (POTENTIAL).
FT DOMAIN 1501 1528 52 (POTENTIAL).
FT TRANSMEM 1529 1556 53 (POTENTIAL).
FT DOMAIN 1557 1584 54 (POTENTIAL).
FT TRANSMEM 1585 1612 55 (POTENTIAL).
FT DOMAIN 1613 1640 56 (POTENTIAL).
FT TRANSMEM 1641 1668 57 (POTENTIAL).
FT DOMAIN 1669 1696 58 (POTENTIAL).
FT TRANSMEM 1697 1724 59 (POTENTIAL).
FT DOMAIN 1725 1752 60 (POTENTIAL).
FT TRANSMEM 1753 1780 61 (POTENTIAL).
FT DOMAIN 1781 1808 62 (POTENTIAL).
FT TRANSMEM 1809 1836 63 (POTENTIAL).
FT DOMAIN 1837 1864 64 (POTENTIAL).
FT TRANSMEM 1865 1892 65 (POTENTIAL).
FT DOMAIN 1893 1920 66 (POTENTIAL).
FT TRANSMEM 1921 1948 67 (POTENTIAL).
FT DOMAIN 1949 1976 68 (POTENTIAL).
FT TRANSMEM 1977 2004 69 (POTENTIAL).
FT DOMAIN 2005 2032 70 (POTENTIAL).
FT TRANSMEM 2033 2060 71 (POTENTIAL).
FT DOMAIN 2061 2088 72 (POTENTIAL).
FT TRANSMEM 2089 2116 73 (POTENTIAL).
FT DOMAIN 2117 2144 74 (POTENTIAL).
FT TRANSMEM 2145 2172 75 (POTENTIAL).
FT DOMAIN 2173 2200 76 (POTENTIAL).
FT TRANSMEM 2201 2228 77 (POTENTIAL).
FT DOMAIN 2229 2256 78 (POTENTIAL).
FT TRANSMEM 2257 2284 79 (POTENTIAL).
FT DOMAIN 2285 2312 80 (POTENTIAL).
FT TRANSMEM 2313 2340 81 (POTENTIAL).
FT DOMAIN 2341 2368 82 (POTENTIAL).
FT TRANSMEM 2369 2396 83 (POTENTIAL).
FT DOMAIN 2397 2424 84 (POTENTIAL).
FT TRANSMEM 2425 2452 85 (POTENTIAL).
FT DOMAIN 2453 2480 86 (POTENTIAL).
FT TRANSMEM 2481 2508 87 (POTENTIAL).
FT DOMAIN 2509 2536 88 (POTENTIAL).
FT TRANSMEM 2537 2564 89 (POTENTIAL).
FT DOMAIN 2565 2592 90 (POTENTIAL).
FT TRANSMEM 2593 2620 91 (POTENTIAL).
FT DOMAIN 2621 2648 92 (POTENTIAL).
FT TRANSMEM 2649 2676 93 (POTENTIAL).
FT DOMAIN 2677 2704 94 (POTENTIAL).
FT TRANSMEM 2705 2732 95 (POTENTIAL).
FT DOMAIN 2733 2760 96 (POTENTIAL).
FT TRANSMEM 2761 2788 97 (POTENTIAL).
FT DOMAIN 2789 2816 98 (POTENTIAL).
FT TRANSMEM 2817 2844 99 (POTENTIAL).
FT DOMAIN 2845 2872 100 (POTENTIAL).
FT TRANSMEM 2873 2900 101 (POTENTIAL).
FT DOMAIN 2901 2928 102 (POTENTIAL).
FT TRANSMEM 2929 2956 103 (POTENTIAL).
FT DOMAIN 2957 2984 104 (POTENTIAL).
FT TRANSMEM 2985 3012 105 (POTENTIAL).
FT DOMAIN 3013 3040 106 (POTENTIAL).
FT TRANSMEM 3041 3068 107 (POTENTIAL).
FT DOMAIN 3069 3096 108 (POTENTIAL).
FT TRANSMEM 3097 3124 109 (POTENTIAL).
FT DOMAIN 3125 3152 110 (POTENTIAL).
FT TRANSMEM 3153 3180 111 (POTENTIAL).
FT DOMAIN 3181 3208 112 (POTENTIAL).
FT TRANSMEM 3209 3236 113 (POTENTIAL).
FT DOMAIN 3237 3264 114 (POTENTIAL).
FT TRANSMEM 3265 3292 115 (POTENTIAL).
FT DOMAIN 3293 3320 116 (POTENTIAL).
FT TRANSMEM 3321 3348 117 (POTENTIAL).
FT DOMAIN 3349 3376 118 (POTENTIAL).
FT TRANSMEM 3377 3404 119 (POTENTIAL).
FT DOMAIN 3405 3432 120 (POTENTIAL).
FT TRANSMEM 3433 3460 121 (POTENTIAL).
FT DOMAIN 3461 3488 122 (POTENTIAL).
FT TRANSMEM 3489 3516 123 (POTENTIAL).
FT DOMAIN 3517 3544 124 (POTENTIAL).
FT TRANSMEM 3545 3572 125 (POTENTIAL).
FT DOMAIN 3573 3600 126 (POTENTIAL).
FT TRANSMEM 3601 3628 127 (POTENTIAL).
FT DOMAIN 3629 3656 128 (POTENTIAL).
FT TRANSMEM 3657 3684 129 (POTENTIAL).
FT DOMAIN 3685 3712 130 (POTENTIAL).
FT TRANSMEM 3713 3740 131 (POTENTIAL).
FT DOMAIN 3741 3768 132 (POTENTIAL).
FT TRANSMEM 3769 3796 133 (POTENTIAL).
FT DOMAIN 3797 3824 134 (POTENTIAL).
FT TRANSMEM 3825 3852 135 (POTENTIAL).
FT DOMAIN 3853 3880 136 (POTENTIAL).
FT TRANSMEM 3881 3908 137 (POTENTIAL).
FT DOMAIN 3909 3936 138 (POTENTIAL).
FT TRANSMEM 3937 3964 139 (POTENTIAL).
FT DOMAIN 3965 3992 140 (POTENTIAL).
FT TRANSMEM 3993 4020 141 (POTENTIAL).
FT DOMAIN 4021 4048 142 (POTENTIAL).
FT TRANSMEM 4049 4076 143 (POTENTIAL).
FT DOMAIN 4077 4104 144 (POTENTIAL).
FT TRANSMEM 4105 4132 145 (POTENTIAL).
FT DOMAIN 4133 4160 146 (POTENTIAL).
FT TRANSMEM 4161 4188 147 (POTENTIAL).
FT DOMAIN 4189 4216 148 (POTENTIAL).
FT TRANSMEM 4217 4244 149 (POTENTIAL).
FT DOMAIN 4245 4272 150 (POTENTIAL).
FT TRANSMEM 4273 4300 151 (POTENTIAL).
FT DOMAIN 4301 4328 152 (POTENTIAL).
FT TRANSMEM 4329 4356 153 (POTENTIAL).
FT DOMAIN 4357 4384 154 (POTENTIAL).
FT TRANSMEM 4385 4412 155 (POTENTIAL).
FT DOMAIN 4413 4440 156 (POTENTIAL).
FT TRANSMEM 4441 4468 157 (POTENTIAL).
FT DOMAIN 4469 4496 158 (POTENTIAL).
FT TRANSMEM 4497 4524 159 (POTENTIAL).
FT DOMAIN 4525 4552 160 (POTENTIAL).
FT TRANSMEM 4553 4580 161 (POTENTIAL).
FT DOMAIN 4581 4608 162 (POTENTIAL).
FT TRANSMEM 4609 4636 163 (POTENTIAL).
FT DOMAIN 4637 4664 164 (POTENTIAL).
FT TRANSMEM 4665 4692 165 (POTENTIAL).
FT DOMAIN 4693 4720 166 (POTENTIAL).
FT TRANSMEM 4721 4748 167 (POTENTIAL).
FT DOMAIN 4749 4776 168 (POTENTIAL).
FT TRANSMEM 4777 4804 169 (POTENTIAL).
FT DOMAIN 4805 4832 170 (POTENTIAL).
FT TRANSMEM 4833 4860 171 (POTENTIAL).
FT DOMAIN 4861 4888 172 (POTENTIAL).
FT TRANSMEM 4889 4916 173 (POTENTIAL).
FT DOMAIN 4917 4944 174 (POTENTIAL).
FT TRANSMEM 4945 4972 175 (POTENTIAL).
FT DOMAIN 4973 5000 176 (POTENTIAL).
FT TRANSMEM 5001 5028 177 (POTENTIAL).
FT DOMAIN 5029 5056 178 (POTENTIAL).
FT TRANSMEM 5057 5084 179 (POTENTIAL).
FT DOMAIN 5085 5112 180 (POTENTIAL).
FT TRANSMEM 5113 5140 181 (POTENTIAL).
FT DOMAIN 5141 5168 182 (POTENTIAL).
FT TRANSMEM 5169 5196 183 (POTENTIAL).
FT DOMAIN 5197 5224 184 (POTENTIAL).
FT TRANSMEM 5225 5252 185 (POTENTIAL).
FT DOMAIN 5253 5280 186 (POTENTIAL).
FT TRANSMEM 5281 5308 187 (POTENTIAL).
FT DOMAIN 5309 5336 188 (POTENTIAL).
FT TRANSMEM 5337 5364 189 (POTENTIAL).
FT DOMAIN 5365 5392 190 (POTENTIAL).
FT TRANSMEM 5393 5420 191 (POTENTIAL).
FT DOMAIN 5421 5448 192 (POTENTIAL).
FT TRANSMEM 5449 5476 193 (POTENTIAL).
FT DOMAIN 5477 5504 194 (POTENTIAL).
FT TRANSMEM 5505 5532 195 (POTENTIAL).
FT DOMAIN 5533 5560 196 (POTENTIAL).
FT TRANSMEM 5561 5588 197 (POTENTIAL).
FT DOMAIN 5589 5616 198 (POTENTIAL).
FT TRANSMEM 5617 5644 199 (POTENTIAL).
FT DOMAIN 5645 5672 200 (POTENTIAL).
FT TRANSMEM 5673 5700 201 (POTENTIAL).
FT DOMAIN 5701 5728 202 (POTENTIAL).
FT TRANSMEM 5729 5756 203 (POTENTIAL).
FT DOMAIN 5757 5784 204 (POTENTIAL).
FT TRANSMEM 5785 5812 205 (POTENTIAL).
FT DOMAIN 5813 5840 206 (POTENTIAL).
FT TRANSMEM 5841 5868 207 (POTENTIAL).
FT DOMAIN 5869 5896 208 (POTENTIAL).
FT TRANSMEM 5897 5924 209 (POTENTIAL).
FT DOMAIN 5925 5952 210 (POTENTIAL).
FT TRANSMEM 5953 5980 211 (POTENTIAL).
FT DOMAIN 5981 6008 212 (POTENTIAL).
FT TRANSMEM 6009 6036 213 (POTENTIAL).
FT DOMAIN 6037 6064 214 (POTENTIAL).
FT TRANSMEM 6065 6092 215 (POTENTIAL).
FT DOMAIN 6093 6120 216 (POTENTIAL).
FT TRANSMEM 6121 6148 217 (POTENTIAL).
FT DOMAIN 6149 6176 218 (POTENTIAL).
FT TRANSMEM 6177 6204 219 (POTENTIAL).
FT DOMAIN 6205 6232 220 (POTENTIAL).
FT TRANSMEM 6233 6260 221 (POTENTIAL).
FT DOMAIN 6261 6288 222 (POTENTIAL).
FT TRANSMEM 6289 6316 223 (POTENTIAL).
FT DOMAIN 6317 6344 224 (POTENTIAL).
FT TRANSMEM 6345 6372 225 (POTENTIAL).
FT DOMAIN 6373 6400 226 (POTENTIAL).
FT TRANSMEM 6401 6428 227 (POTENTIAL).
FT DOMAIN 6429 6456 228 (POTENTIAL).
FT TRANSMEM 6457 6484 229 (POTENTIAL).
FT DOMAIN 6485 6512 230 (POTENTIAL).
FT TRANSMEM 6513 6540 231 (POTENTIAL).
FT DOMAIN 6541 6568 232 (POTENTIAL).
FT TRANSMEM 6569 6596 233 (POTENTIAL).
FT DOMAIN 6597 6624 234 (POTENTIAL).
FT TRANSMEM 6625 6652 235 (POTENTIAL).
FT DOMAIN 6653 6680 236 (POTENTIAL).
FT TRANSMEM 6681 6708 237 (POTENTIAL).
FT DOMAIN 6709 6736 238 (POTENTIAL).
FT TRANSMEM 6737 6764 239 (POTENTIAL).
FT DOMAIN 6765 6792 240 (POTENTIAL).
FT TRANSMEM 6793 6820 241 (POTENTIAL).
FT DOMAIN 6821 6848 242 (POTENTIAL).
FT TRANSMEM 6849 6876 243 (POTENTIAL).
FT DOMAIN 6877 6904 244 (POTENTIAL).
FT TRANSMEM 6905 6932 245 (POTENTIAL).
FT DOMAIN 6933 6960 246 (POTENTIAL).
FT TRANSMEM 6961 6988 247 (POTENTIAL).
FT DOMAIN 6989 7016 248 (POTENTIAL).
FT TRANSMEM 7017 7044 249 (POTENTIAL).
FT DOMAIN 7045 7072 250 (POTENTIAL).
FT TRANSMEM 7073 7100 251 (POTENTIAL).
FT DOMAIN 7101 7128 252 (POTENTIAL).
FT TRANSMEM 7129 7156 253 (POTENTIAL).
FT DOMAIN 7157 7184 254 (POTENTIAL).
FT TRANSMEM 7185 7212 255 (POTENTIAL).
FT DOMAIN 7213 7240 256 (POTENTIAL).
FT TRANSMEM 7241 7268 257 (POTENTIAL).
FT DOMAIN 7269 7296 258 (POTENTIAL).
FT TRANSMEM 7297 7324 259 (POTENTIAL).
FT DOMAIN 7325 7352 260 (POTENTIAL).
FT TRANSMEM 7353 7380 261 (POTENTIAL).
FT DOMAIN 7381 7408 262 (POTENTIAL).
FT TRANSMEM 7409 7436 263 (POTENTIAL).
FT DOMAIN 7437 7464 264 (POTENTIAL).
FT TRANSMEM 7465 7492 265 (POTENTIAL).
FT DOMAIN 7493 7520 266 (POTENTIAL).
FT TRANSMEM 7521 7548 267 (POTENTIAL).
FT DOMAIN 7549 7576 268 (POTENTIAL).
FT TRANSMEM 7577 7604 269 (POTENTIAL).
FT DOMAIN 7605 7632 270 (POTENTIAL).
FT TRANSMEM 7633 7660 271 (POTENTIAL).
FT DOMAIN 7661 7688 272 (POTENTIAL).
FT TRANSMEM 7689 7716 273 (POTENTIAL).
FT DOMAIN 7717 7744 274 (POTENTIAL).
FT TRANSMEM 7745 7772 275 (POTENTIAL).
FT DOMAIN 7773 7800 276 (POTENTIAL).
FT TRANSMEM 7801 7828 277 (POTENTIAL).
FT DOMAIN 7829 7856 278 (POTENTIAL).
FT TRANSMEM 7857 7884 279 (POTENTIAL).
FT DOMAIN 7885 7912 280 (POTENTIAL).
FT TRANSMEM 7913 7940 281 (POTENTIAL).
FT DOMAIN 7941 7968 282 (POTENTIAL).
FT TRANSMEM 7969 7996 283 (POTENTIAL).
FT DOMAIN 7997 8024 284 (POTENTIAL).
FT TRANSMEM 8025 8052 285 (POTENTIAL).
FT DOMAIN 8053 8080 286 (POTENTIAL).
FT TRANSMEM 8081 8108 287 (POTENTIAL).
FT DOMAIN 8109 8136 288 (POTENTIAL).
FT TRANSMEM 8137 8164 289 (POTENTIAL).
FT DOMAIN 8165 8192 290 (POTENTIAL).
FT TRANSMEM 8193 8220 291 (POTENTIAL).
FT DOMAIN 8221 8248 292 (POTENTIAL).
FT TRANSMEM 8249 8276 293 (POTENTIAL).
FT DOMAIN 8277 8304 294 (POTENTIAL).
FT TRANSMEM 8305 8332 295 (POTENTIAL).
FT DOMAIN 8333 8360 296 (POTENTIAL).
FT TRANSMEM 8361 8388 297 (POTENTIAL).
FT DOMAIN 8389 8416 298 (POTENTIAL).
FT TRANSMEM 8417 8444 299 (POTENTIAL).
FT DOMAIN 8445 8472 300 (POTENTIAL).
FT TRANSMEM 8473 8500 301 (POTENTIAL).
FT DOMAIN 8501 8528 302 (POTENTIAL).
FT TRANSMEM 8529 8556 303 (POTENTIAL).
FT DOMAIN 8557 8584 304 (POTENTIAL).
FT TRANSMEM 8585 8612 305 (POTENTIAL).
FT DOMAIN 8613 8640 306 (POTENTIAL).
FT TRANSMEM 8641 8668 307 (POTENTIAL).
FT DOMAIN 8669 8696 308 (POTENTIAL).
FT TRANSMEM 8697 8724 309 (POTENTIAL).
FT DOMAIN 8725 8752 310 (POTENTIAL).
FT TRANSMEM 8753 8780 311 (POTENTIAL).
FT DOMAIN 8781 8808 312 (POTENTIAL).
FT TRANSMEM 8809 8836 313 (POTENTIAL).
FT DOMAIN 8837 8864 314 (POTENTIAL).
FT TRANSMEM 8865 8892 315 (POTENTIAL).
FT DOMAIN 8893 8920 316 (POTENTIAL).
FT TRANSMEM 8921 8948 317 (POTENTIAL).
FT DOMAIN 8949 8976 318 (POTENTIAL).
FT TRANSMEM 8977 9004 319 (POTENTIAL).
FT DOMAIN 9005 9032 320 (POTENTIAL).
FT TRANSMEM 9033 9060 321 (POTENTIAL).
FT DOMAIN 9061 9088 322 (POTENTIAL).
FT TRANSMEM 9089 9116 323 (POTENTIAL).
FT DOMAIN 9117 9144 324 (POTENTIAL).
FT TRANSMEM 9145 9172 325 (POTENTIAL).
FT DOMAIN 9173 9200 326 (POTENTIAL).
FT TRANSMEM 9201 9228 327 (POTENTIAL).
FT DOMAIN 9229 9256 328 (POTENTIAL).
FT TRANSMEM 9257 9284 329 (POTENTIAL).
FT DOMAIN 9285 9312 330 (POTENTIAL).
FT TRANSMEM 9313 9340 331 (POTENTIAL).
FT DOMAIN 9341 9368 332 (POTENTIAL).
FT TRANSMEM 9369 9396 333 (POTENTIAL).
FT DOMAIN 9397 9424 334 (POTENTIAL).
FT TRANSMEM 9425 9452 335 (POTENTIAL).
FT DOMAIN 9453 9480 336 (POTENTIAL).
FT TRANSMEM 9481 9508 337 (POTENTIAL).
FT DOMAIN 9509 9536 338 (POTENTIAL).
FT TRANSMEM 9537 9564 339 (POTENTIAL).
FT DOMAIN 9565 9592 340 (POTENTIAL).
FT TRANSMEM 9593 9620 341 (POTENTIAL).
FT DOMAIN 9621 9648 342 (POTENTIAL).
FT TRANSMEM 9649 9676 343 (POTENTIAL).
FT DOMAIN 9677 9704 344 (POTENTIAL).
FT TRANSMEM 9705 9732 345 (POTENTIAL).
FT DOMAIN 9733 9760 346 (POTENTIAL).
FT TRANSMEM 9761 9788 347 (POTENTIAL).
FT DOMAIN 9789 9816 348 (POTENTIAL).
FT TRANSMEM 9817 9844 349 (POTENTIAL).
FT DOMAIN 9845 9872 350 (POTENTIAL).
FT TRANSMEM 9873 9900 351 (POTENTIAL).
FT DOMAIN 9901 9928 352 (POTENTIAL).
FT TRANSMEM 9929 9956 353 (POTENTIAL).
FT DOMAIN 9957 9984 354 (POTENTIAL).
FT TRANSMEM 9985 10000 355 (POTENTIAL).

```

Db 42 LSYQIITSLFGLIILCSIFGNSCVAAIALER--SLQNVANYLIGSLAVTDLNVSVLV 99  
QY 86 PMWATSVPLFMPINSHFCTALVSLTHLFAFASVNTIVASVDRYLSIHPLVSPMTQ 145  
Db 100 PMALVLYLKNKWTIGQIDICFLALDYVLCCTSSILHICAILDYMATIDPIDYVNRTP 159  
QY 146 RRGVLLLYGTWIVAILIOTPPYLGWGOAFDERNALCSMTWGSAPSSTILSVSFYIPL 205  
Db 160 RRAAVLSVWVLIGFSTISIPMLGW-RSAEDBRAPDACIIT-SQDPGTYISTFCAFYIPL 217  
QY 206 IWMVACTSVVFCARRRQHALLYNKRHSLEYRKDCVENDESGAEKKEEFQDESRRQ 265  
Db 218 ILMVLYGRIFPKAR-----FRIRKTYKTEKAKASVDMCLTILSPAVPHKR 262  
QY 266 HEGEVKAKEGMEAKDGLAKKEGSTSTSESVARSGSEVEBRESSTVYASDGSMEKKGST 325  
Db 263 ANGA-----VSAEMKRGYKFKPSPCANCAVRHGE-- 294  
QY 326 KVEENSMKADKGRTEVNCSDIDGEDMERGEDDINFESEDDVEAVNIPESLPPSRNSNS 385  
Db 295 -----MESLEIEVNSN-SKTHLPLPNTPOG--SSHENINE 327  
QY 386 NPLPLP-----CYOCKAKAKYFIITISYVLSLGPYCLAVLAVWDVETQVQWVITII 440  
Db 328 KTGTFERKIALAREKRTVKTIGIMGTIFCWLPEFIVALTLPFCAENCYMPEWLGAVIN 387  
QY 441 WLFELCCIHRYVYGYMHTTKREIOMLKKFPCK 475  
Db 388 WLGYSNLSNLNPIYAFNKPDSAFKKILR--CK 419  
RESULT 10  
ALIA\_MOUSE STANDARD; PRT: 466 AA.  
AC P97718; 054913;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Alpha-1a adrenergic receptor (Alpha 1a-adrenoceptor) (Alpha-1c  
adrenoceptor).  
GN ADRA1A OR ADRA1C.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. AND CHARACTERIZATION.  
RC STRAIN=CD-1; TISSUE=Brain, and Kidney;  
RX MEDLINE=98292316; Pubmed=9630362;  
RA Xiao L., Scofield M.A., Jeffries W.B.;  
RT "Molecular cloning, expression and characterization of cDNA encoding a  
mouse alpha1a-adrenoceptor".  
RL Br. J. Pharmacol. 124:213-221(1998).  
RN [2]  
RP SEQUENCE OF 197-280 FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=96064818; Pubmed=7595531;  
RA Alonso-Llamazares A., Zamanillo D., Casanova E., Ovalle S., Calvo P.,  
RA Chinchetru M.A.;  
RT "Molecular cloning of alpha 1d-adrenergic receptor and tissue  
distribution of three alpha 1-adrenergic receptor subtypes in  
mouse".  
RL J. Neurochem. 65:2387-2392(1995).  
CC - FUNCTION: THIS ALPHA-ADRENERGIC RECEPTOR MEDIATES ITS ACTION BY  
ASSOCIATION WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-  
CALCIUM SECOND MESSENGER SYSTEM. ITS EFFECT IS MEDIATED BY G(O)  
AND G(11) PROTEINS.  
CC - SUBCELLULAR LOCATION: Integral membrane protein.  
CC - PTM: CARBOXYL-TERMINAL SER OR THR RESIDUES MAY BE PHOSPHORYLATED  
(BY SIMILARITY).  
CC - SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF031431; AAC02658.1; -  
DR EMBL: S80220; AAB47044.1; ALT\_INT.  
DR GDB: MGI:104773; Agra1a.  
DR InterPro: IPR00276; GPCR\_Rhodpsn.  
DR Pfam: PF00001; 7tm\_1; 1.  
DR PRINTS: PR00237; GPCR\_RHODPSN.  
DR PROSITE: PS00237; G-PROTEIN\_RECP\_F1\_1; 1.  
DR PROSITE: PS50262; G-PROTEIN\_RECP\_F1\_2; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
KW Multigene family; Phosphorylation; Lipoprotein; Palmitate.  
FT DOMAIN 1 27  
FT TRANSSEM 28 51  
FT DOMAIN 52 64  
FT TRANSSEM 65 88  
FT DOMAIN 89 99  
FT TRANSSEM 100 122  
FT DOMAIN 123 143  
FT TRANSSEM 144 167  
FT DOMAIN 168 181  
FT TRANSSEM 182 205  
FT DOMAIN 206 273  
FT TRANSSEM 274 297  
FT DOMAIN 298 305  
FT TRANSSEM 306 329  
FT DOMAIN 330 466  
FT CARBOHYD 7 7  
FT CARBOHYD 13 13  
FT CARBOHYD 22 22  
FT MOD\_RES 215 215  
FT DISULFID 99 176  
FT LIPID 345 345  
FT SEQUENCE 466 AA; 51762 MW; 36495A53DBD34DA CRC64;  
SQ  
Query Match 12.6%; Score 333; DB 1; Length 466;  
Best Local Similarity 22.5%; Pred. No. 2.1e-12;  
Matches 114; Conservative 75; Mismatches 184; Indels 134; Gaps 13;  
QY 7 NSTRENSHTCMPLSKMPISLANGIIRTYLVIFPLASFGNIVLALVLRKQOLLOYT 66  
Db 7 NASGGSNCTH---PPAONISKRA--ILGVLIGGLIFGVGLNVLVLSVCHRHLSHYT 61  
QY 67 NRETFNLVTDLLOISLVA PWVATSVPLFMPINSHFCTALVSLTHLFAFASVNTIVLVS 126  
Db 62 HYIVNLAVALDLTSTYLPFSALFELIGYAFGRVFCNIMAAVDVLCCTASIMGLCTIS 121  
QY 127 VDRYLSIHLPLSPSKMTQRGYLLYTWTVAIILOSTPPYLGWGOAFDERNALCSMIW 186  
Db 122 IDRYIGVSYPLRYPTIYQIRGVALLCVMALSIYISIGPLFGWMOQ-PEDETICQT-- 178  
QY 187 GASPSYTTLSVSVFTVPLIYMIACYSVFCARRQHALLYNKRHSLEYVAKCVENED 246  
Db 179 NEEPGYVLFSAIGSRYPVLTJIIWYCRVYVARR-----ESRG 217  
QY 247 EGAEKKEEFQDESFRROHGEVAKGRMEAKDGLAKKEGSTNGTSESVEARGSBEV 306  
Db 218 LKSGIKTDKSDSEOVYTLRIHKKNVA-----EGS-GVSSA----- 251  
QY 307 RESSTVASDGSMEKKEGSTKVEENSMKADKGRTEVNCSDIDGEDMERGEDDINFESEDD 366  
Db 252 -----KNKT-----HFSVRLKFSREX 268  
QY 367 VEAVNIPESLPPSRNSNPNPLPRCYOCKAKAKYFIITISYVLSLGPYCLAVLAVWD 426  
Db 269 -----KAAKTLGIYVGCFLCWLPLF-FLVWVIGSFF 298  
QY 427 VETQVQWVITIIWLFELCCIHRYVYGYMHTTKREIOMLKKFPCKEKPREDSD--- 483

DR PRINTS: PR00237; GPCRHRHODPSN.  
 DR PROSITE: PS00237; G-PROTEIN RECEPTOR.F1.1; 1.  
 DR PROSITE: PS0262; G-PROTEIN RECEPTOR.F1.2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 MultiGene family.

FT	DOMAIN	1	36	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	37	62	1 (POTENTIAL).
FT	DOMAIN	63	73	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	74	98	2 (POTENTIAL).
FT	DOMAIN	99	109	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	110	132	3 (POTENTIAL).
FT	DOMAIN	133	152	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	153	178	4 (POTENTIAL).
FT	DOMAIN	179	191	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	192	217	5 (POTENTIAL).
FT	DOMAIN	218	345	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	346	367	6 (POTENTIAL).
FT	DOMAIN	368	378	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	379	403	7 (POTENTIAL).
FT	DOMAIN	404	421	CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	10	10	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	11	11	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	24	24	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	30	30	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	109	187	BY SIMILARITY.
FT	VARIANT	17	19	EPR -> DHL (IN STRAIN C3H/AN).
FT	CONFLICT	177	177	A -> T (IN REF. 3).
FT	CONFLICT	242	242	F -> L (IN REF. 3).
FT	CONFLICT	247	247	A -> V (IN REF. 3).
FT	CONFLICT	263	263	C -> W (IN REF. 3).
FT	CONFLICT	304	304	D -> H (IN REF. 3).
SO	SEQUENCE	421 AA;	46123 MW;	CD8519D37A1F49B CRC64;

Query Match 12.7%; Score 335.5; DB 1; Length 421;  
 Best Local Similarity 23.2%; Pred. No. 1.4e-12;  
 Matches 106; Conservative 88; Mismatches 190; Indels 73; Gaps 12;

QY	26	ISLHGIRSTVIVFLASFGVGN--IVLALVLRKPOLQVTRFIFNLVTDLQISL	83
DB	31	VTFSTYQVITSLGLTGLFCALVAGNACVVAIALER--SLQNVANVILGLSLAVDLMVSVL	88
QY	84	VAPVVAATVPLFMPVPLNSHETALVSLTHLFAFASVNTIVLVSVDRYLSIIHPLSPYSKM	143
DB	89	VLPMAALVQLNKTWTLQGVTCDFIALDVLCSTSIHLCAIALDRWATPDPIDVYKRR	148
QY	144	TORGYLLGTMTVALQSTPRPLVGGGAQAFDERNA--LCSMIGASPSYTIISVSEFIV	202
DB	149	TPRAAALISLTLGLFLISIPMLGW--RAVEDRSNNECT--SKDHGYTISTFGAFY	205
QY	203	IPLVIMACYSVFCARROHALLYNKRSLEVRKVCVENEDDEGAKEEFQDESEF	262
DB	206	IPLMLLVILGRIFRAR-----FRIRKYKVKVKKCA-----	238
QY	263	ROHEGEVNAKEGMEAKDGLAKESGTSTSESVARGSEVEESSIVASDGM--EGK	321
DB	239	-----GTSFGTSAPPRKSLNG--QPGSGDCRSRAENRAVGTPCANAVRQGE	285
QY	322	EGST--KVENSMKADKGRVENQCSIDLGDHMEFGEDDINSEDDVEANINPESP--P	378
DB	286	DDALEVIEVHRGNSG-----DLPDSEGAASYVACLERK	324
QY	379	SRNSNSNPPLPRCYCKAAKAVIFIIIFSVYLSLGPYCFGLAVLAVWVDETQVQWYTI	438
DB	325	NERAEAKRMKALAREKRYKTGIIINGTFLCMLPFFIYALVLPDSSCHMPELIGAI	384
QY	439	ITLFFLOCCIHPRVYGYMKTIKKETQDMKKPFCK	475
DB	385	INMLGYSNLNPVYAYVFNKDFONAFKIKCKFCR	421

RESULT 9  
 SHIA\_FUGRU STANDARD; PRT; 423 AA.

AC 042385;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE 5-hydroxytryptamine 1A-alpha receptor (5-HT<sub>1A</sub>-alpha) (Serotonin  
 DE receptor) (5-HT<sub>1A</sub>-alpha) (F1A).  
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;  
 OC Tetraodontidae; Takifugu.  
 OX NCBI\_TaxID=31033;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=97361762; PubMed=9218723;  
 RA Yamaguchi F., Brenner S.;  
 RT Molecular cloning of 5-hydroxytryptamine (5-HT) type 1 receptor  
 RT genes from the Japanese puffer fish, Fugu rubripes.";  
 RL Gene 191:219-223(1997).  
 CC -1- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR 5-  
 CC HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC HORMONE THAT FUNCTIONS  
 CC AS A NEUROTRANSMITTER, A HORMONE, AND A MITOGEN. THE ACTIVITY OF  
 CC THIS RECEPTOR IS MEDIATED BY G-PROTEINS THAT INHIBITS ADENYLYLATE  
 CC CYCLASE ACTIVITY (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC STRONGEST TO THE OTHER 5HT-1 SUBTYPE RECEPTORS.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X95936; CAA65175.1; -.  
 DR GCRD: GCR\_2429; -.  
 DR InterPro: IPR00276; GPCR\_Rhodopsn.  
 DR Pfam: PF00001; 7tm\_1; 1.  
 DR PRINTS: PR00237; GPCRHRHODPSN.  
 DR PROSITE: PS00237; G-PROTEIN RECEPTOR.F1.1; 1.  
 DR PROSITE: PS0262; G-PROTEIN RECEPTOR.F1.2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 MultiGene family.

FT	DOMAIN	1	45	EXTRACELLULAR (POTENTIAL).
FT <td>TRANSMEM</td> <td>46</td> <td>71</td> <td>1 (POTENTIAL).</td>	TRANSMEM	46	71	1 (POTENTIAL).
FT <td>DOMAIN</td> <td>72</td> <td>82</td> <td>CYTOPLASMIC (POTENTIAL).</td>	DOMAIN	72	82	CYTOPLASMIC (POTENTIAL).
FT <td>TRANSMEM</td> <td>83</td> <td>107</td> <td>2 (POTENTIAL).</td>	TRANSMEM	83	107	2 (POTENTIAL).
FT <td>DOMAIN</td> <td>108</td> <td>118</td> <td>EXTRACELLULAR (POTENTIAL).</td>	DOMAIN	108	118	EXTRACELLULAR (POTENTIAL).
FT <td>TRANSMEM</td> <td>119</td> <td>141</td> <td>3 (POTENTIAL).</td>	TRANSMEM	119	141	3 (POTENTIAL).
FT <td>DOMAIN</td> <td>142</td> <td>161</td> <td>CYTOPLASMIC (POTENTIAL).</td>	DOMAIN	142	161	CYTOPLASMIC (POTENTIAL).
FT <td>TRANSMEM</td> <td>162</td> <td>186</td> <td>4 (POTENTIAL).</td>	TRANSMEM	162	186	4 (POTENTIAL).
FT <td>DOMAIN</td> <td>188</td> <td>200</td> <td>EXTRACELLULAR (POTENTIAL).</td>	DOMAIN	188	200	EXTRACELLULAR (POTENTIAL).
FT <td>TRANSMEM</td> <td>201</td> <td>226</td> <td>5 (POTENTIAL).</td>	TRANSMEM	201	226	5 (POTENTIAL).
FT <td>DOMAIN</td> <td>227</td> <td>346</td> <td>CYTOPLASMIC (POTENTIAL).</td>	DOMAIN	227	346	CYTOPLASMIC (POTENTIAL).
FT <td>TRANSMEM</td> <td>347</td> <td>368</td> <td>6 (POTENTIAL).</td>	TRANSMEM	347	368	6 (POTENTIAL).
FT <td>DOMAIN</td> <td>369</td> <td>379</td> <td>EXTRACELLULAR (POTENTIAL).</td>	DOMAIN	369	379	EXTRACELLULAR (POTENTIAL).
FT <td>TRANSMEM</td> <td>380</td> <td>404</td> <td>7 (POTENTIAL).</td>	TRANSMEM	380	404	7 (POTENTIAL).
FT <td>DOMAIN</td> <td>405</td> <td>423</td> <td>CYTOPLASMIC (POTENTIAL).</td>	DOMAIN	405	423	CYTOPLASMIC (POTENTIAL).
FT <td>DISULFID</td> <td>118</td> <td>196</td> <td>BY SIMILARITY.</td>	DISULFID	118	196	BY SIMILARITY.
FT <td>CARBOHYD</td> <td>9</td> <td>9</td> <td>N-LINKED (GLCNAC. . .) (POTENTIAL).</td>	CARBOHYD	9	9	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>12</td> <td>12</td> <td>N-LINKED (GLCNAC. . .) (POTENTIAL).</td>	CARBOHYD	12	12	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>30</td> <td>30</td> <td>N-LINKED (GLCNAC. . .) (POTENTIAL).</td>	CARBOHYD	30	30	N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	423 AA;	47000 MW;	7B130862B840190F CRC64;

Query Match 12.6%; Score 333; DB 1; Length 423;  
 Best Local Similarity 24.0%; Pred. No. 1.9e-12;  
 Matches 109; Conservative 74; Mismatches 188; Indels 84; Gaps 11;







```

RC      TISSUE-Prostate; PubMed=8396931;
RX      MEDLINE=93384619;
RA      Hirasawa A., Horie K., Tanaka T., Takagaki K., Mural M., Yano J.,
RA      Tsubimoto G.;
RT      "Cloning, functional expression and tissue distribution of human CDNA
RT      for the alpha 1C-adrenergic receptor.";
RL      Biochem. Biophys. Res. Commun. 195;902-909(1993).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Tseng-Crank J.C., Goetz A., Saussy D., Robertson K.M., Hazum S.,
RA      Hatzilip J., Godinot N., Wisely B., Robertson C.N., Kost T.;
RL      Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RX      TISSUE=Lymphocytes; and Hippocampus;
RA      MEDLINE=94239386; PubMed=8183249;
RA      Porray C., Bard J.A., Wetzel J.M., Chiu G., Shapiro E., Tang R.,
RA      Lepore H., Hartig P.R., Weinschenk R.L., Branchek T.A.,
RA      Gluchowski C.;
RT      "The alpha 1-adrenergic receptor that mediates smooth muscle
RT      contraction in human prostate has the pharmacological properties of
RT      the cloned human alpha 1c subtype.";
RL      Mol. Pharmacol. 45;703-706(1994).
RN      [4]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=95114877; PubMed=7815325;
RA      Schain D.A., Johnston G.I., Page S.O., Mosley M.J., Wilson K.H.,
RA      Norman N.P., Campbell S., Fidock M.D., Furness L.M.,
RA      Parry-Smith D.J., Peter B., Bailey D.S.;
RT      "Cloning and pharmacological characterization of human alpha-1
RT      adrenergic receptors: sequence corrections and direct comparison with
RT      other species homologues.";
RL      J. Pharmacol. Exp. Ther. 272;134-142(1995).
CC      -1- FUNCTION: THIS ALPHA-ADRENERGIC RECEPTOR MEDIATES ITS ACTION BY
CC      ASSOCIATION WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC      CALCIUM SECOND MESSENGER SYSTEM. ITS EFFECT IS MEDIATED BY G(O)
CC      AND G(11) PROTEINS.
CC      -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC      -1- TISSUE SPECIFICITY: HEART, BRAIN, LIVER AND PROSTATE, BUT NOT IN
CC      KIDNEY, LUNG, ADRENAL, AORTA AND PITUITARY.
CC      -1- PTM: CARBOXYL-TERMINAL SER OR THR RESIDUES MAY BE PHOSPHORYLATED
CC      (BY SIMILARITY).
CC      -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; D25235; BAA04960.1; -
DR      EMBL; U02569; AAA93114.1; -
DR      EMBL; U03866; AAB60353.1; -
DR      EMBL; L31774; AAB59486.1; -
DR      GCRDB; GCR_0542; -
DR      GCRDB; GCR_0802; -
DR      GCRDB; GCR_1016; -
DR      GCRDB; GCR_1881; -
DR      MIM; 104221; -
DR      InterPro; IPR000276; GPCR_Rhodpsn.
DR      Pfam; PF00001; 7tm.1; 1.
DR      PRINTS; PR00237; GPCRHHODPSN.
DR      PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR      PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW      G-protein coupled receptor; Transmembrane; Glycoprotein;
KW      Multigene family; Phosphorylation; Lipoprotein; Palmitate.
FT      TRANSMEM 1 27
FT      DOMAIN 28 51
FT      TRANSMEM 52 64
FT      TRANSMEM 65 88
FT      TRANSMEM 89 99
FT      DOMAIN 89 99
EXTRACELLULAR (POTENTIAL).
CYTOPLASMIC (POTENTIAL).
2 (POTENTIAL).
EXTRACELLULAR (POTENTIAL).

```

Query Match	Best Local Similarity	Score	DB 1:	Length	466:
Matches	108; Conservative	12.8%; Pred. No. 16-12;	80; Mismatches 187;	Indels 131;	Gaps 11
Qy	8	STRRESNHTCMPLSMPLSMPLSLAHGIRTSYLVIFFLAASFEVGNIVLVLQRRKPOLLOVYN	67		
Db	5	SGNADSSNCTOP--PAPNISKAILLGVILGGLIFLGVGNITVLVLSVACHRHLSYTH	62		
Qy	68	REFIENLVTDLLQISLVADVVVATVSPLEMPPLNSHFCITALVSLTHAFASVNTVLVGY	127		
Db	63	YIVNLAVALDLITFVLFPFSALEFVLGVAFGVRCNIMAAVDVLCTGASINGLCTIST	122		
Qy	128	DRLSTIHPFLSPYSKKTORGVLVLYGTVMVALLOSTPLPYGGAAFDERNALGSMING	187		
Db	123	DRIGVSYPLRITYTQIRGLMALLCVMAISLVISIGLFGRRQRA-PEDETICQI--N	179		
Qy	188	ASPSYITLVSFIVPIPLVMACYSVFCARROHALLYNKRHSLEVYKQVENEDE	247		
Db	180	EEPGVYLFSAISGFYPLAILVMYCRVYVAKR-----ESRGL	218		
Qy	248	EGAEKKEEPDDESEFPROHGEVAKEGRMADKSLAKEGSGTGSSEVSSEARSGSEVR	307		
Db	219	KSLGKTDKSDSEQYTLRIHKNAFA-----GSGMASATKTHFSVRL-	261		
Qy	308	ESSTVASDGSMEKEGSTKVEENSMRADKGRTEVYNOCSIDLGEDDMFEGEDDINFSEDDV	367		
Db	262	-----LKFSREK-	268		
Qy	368	EAVNIPESLPPSRNSNSNPRLPRCYOCKAKAVYFIILFSYVLSLGYCFELAVLAVWVDV	427		
Db	269	-----KAKTIGIYVGCFLVCMLPF-FLVMYIGSFPF	299		
Qy	428	ETGQPOWVITIIILFFLOCCIHPRVYVYVYHKTIKKEIDMLK-KFCECKEPKE-----D	482		
Db	300	DFKSEIVEKFIKIVFLGYLNSCINIIYPCSGQEFKKAFQVNLRIQCLCRKSSKHALGYT	359		
Qy	483	SHDLPCTEGTEGTEGKI--VPSTDSATF	507		
Db	360	LHPPSQAVEGQHKDMVRIPVGSRETF	385		
RESULT 6					
ALIA_CAVPO					
ID	ALIA_CAVPO	STANDARD:	PRT:	466	AA.
AC	OSWU25;				
DT	16-OCT-2001 (Rel. 40, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Alpha-1A adrenergic receptor (Alpha 1A-adrenoceptor) (Alpha-1C				

```

Db 122 IDRIQSVPLRYPTIYVORGLRALLCVAFSLVSGPLFGMRQAPDD-ETICQI-- 178
QY 187 GASPSYILSVSEFYIPLVIMTACYSVFCARROHALLYNKRHSLEVVKOCVENED 246
Db 179 NEEPGYLSFALSIFYPLIITIAMCRVYVAKR-----ESRG 217
QY 247 EGAEEKKEEPQDSEFFRRQHEGEVKAKEGMEAKDGLAKEGSTGTSTSSSEVARGSEEV 306
Db 218 LKGLTKDKSDSQVTLRIHKRNAP-----GSSGVASA----- 251
QY 307 RESSTVASDGSMEGKEBSTVYENSKADGRTVEVNOCSIDLGEDMEFEGEDDINSEDD 366
Db 252 -----KNKT-----HFSVRLKTFESREK 268
QY 367 VEAVNIPESLPSPRRNSNPPLPRCYCKAKAVIFIIIFSVLSGPGCFVLAVWVD 426
Db 269 -----KAAKTLGIVGCFVLCWLPF-FLVWPIGTSFF 298
QY 427 VETQVQWVITIIILWLEFLOCCIHPPVGYMHKTIKEIDMK 470
Db 299 PDKRPETVFKIVFWLGYLNSCINPIIYPCSSQEFKKAQNVAK 342

RESULT 4
ALIA_ORYLA
ID ALIA_ORYLA STANDARD; PRT; 470 AA.
AC 091175;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Alpha-1A adrenergic receptor (MARL).
OS Oryzias latipes (Medaka fish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
  Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
  Belontiiformes; Adrianichthyidae; Oryziatidae; Oryzias.
NCBI_TaxID=8090;
RP [1]
  SEQUENCE FROM N.A.
RX MEDLINE=96184522; PubMed=8654394;
  Yasuoka A., Abe K., Arai S., Emori Y.;
  "Molecular cloning and functional expression of the
  alpha1A-adrenoceptor of Medaka fish, Oryzias latipes.";
  Eur. J. Biochem. 235:501-507(1996).
RL
CC -1- FUNCTION: THIS ALPHA-ADRENERGIC RECEPTOR MEDIATES ITS ACTION BY
  ASSOCIATION WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
  CALCIUM SECOND MESSENGER SYSTEM (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL, D63859; BAA09921.1; -
DR GCRD; GCR_1567; -
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECP_FL_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
  Multigene family; Phosphorylation; Lipoprotein; Palmitate.
KW DOMAIN 1 27 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 28 51 1 (POTENTIAL).
FT DOMAIN 52 64 1 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 65 88 2 (POTENTIAL).
FT DOMAIN 89 99 3 (POTENTIAL).
FT TRANSMEM 100 122 3 (POTENTIAL).
RP

```

```

FT DOMAIN 123 143 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 144 167 4 (POTENTIAL).
FT DOMAIN 168 181 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 182 205 5 (POTENTIAL).
FT DOMAIN 206 221 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 222 295 6 (POTENTIAL).
FT DOMAIN 296 303 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 304 327 7 (POTENTIAL).
FT DOMAIN 328 470 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 9 9 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 12 12 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 99 176 BY SIMILARITY.
FT LIPID 343 343 PALMITATE (POTENTIAL).
FT DOMAIN 351 354 POLY-HIS.
SQ SEQUENCE 470 AA; 51925 MW; D4F7A8303061D4E CRC64;

Query Match 12.98; Score 340; DB 1; Length 470;
Best Local Similarity 23.88; Pred. No. 8,6e-13;
Matches 119; Conservative 66; Mismatches 180; Indels 134; Gaps 14;

QY 12 SNSGHTCMPLSKMPDISLAHGIRSTVLVIFLASFGVGNIVLALYQKRPOLLQVYTNREIF 71
Db 11 SNCSHVLP-----ELNIVKAVLVGMVLGIFLFGVIGNILVLSVCHRHQLQVITYFIV 66
QY 72 NLVTDLLQISLVAPWVAVTSVPLPMPNSHCTALVSLTHLFASVNTIVLVSDRYL 131
Db 67 NLAVADLLSTVLVPSFAIFEILDRAWGEGVFCIMMAVAVDLCTASIMSLCVISVDRI 126
QY 132 SIHPLSPSKMTORRGVLLGYTVAIIIOSTPPLYGMOGAARDERNALCSMTWGSPTS 191
Db 127 GVSYPLEKYPALMTKRRALLVLMVLSVSIISGLPGWKQKPA--PEDEYCKIT--BEPG 183
QY 192 YTLISVSEFYIPLVIMTACYSVFCARROHALLYNKRHSLEVVKDCVENEDEGAE 251
Db 184 YATISAVGSFTPLAITIAMCRVYVAAQKSRK-----KKG-Q 222
QY 252 KKEEFQDSEFFRRQHEGEVKAKEGMEAKDGLAKEGSTGTSTSSSEVARGSEEVRESST 311
Db 223 KIEKSDSEQVILRMHRG-----NTT 242
QY 312 VASGSMGKEGSTKVEBNSKKADGRTVEVNOCSIDLGEDMEFEGEDDINSEDDVAVN 371
Db 243 VSED-----EALRSRT-----HFAIRLKLFSREK----- 266
QY 372 IPESLPSPRRNSNPPLPRCYCKAKAVIFIIIFSVLSGPGCFVLAVWVDVEYQ 431
Db 267 -----KAAKTLGIVGCFVLCWLP--FLVLPISIPAYR 300
QY 432 P-QWVITIIILWLEFLOCCIHPPVGYMHKTIKEIDMKKFKCKEPPKEDSHPDLPCT 490
Db 301 PSDVFPKTFITFWLGYFNSCINPIIILCSNQEKKAFQSL--GVHCLRMTPRAHHH-HLSVG 358
QY 491 EGGTEG-KIVDSYDSTAF 508
Db 359 QSQTOGHSLRTISDSKGP 377

RESULT 5
ALIA_HUMAN
ID ALIA_HUMAN STANDARD; PRT; 466 AA.
AC P35348;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-1A adrenergic receptor (Alpha 1A-adrenoceptor) (Alpha-1C
  adrenergic receptor).
GN ADRA1A OR ADRA1C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RP [1]
  SEQUENCE FROM N.A.

```

```

FT TRANSMEM 379 403 7 (POTENTIAL).
FT DOMAIN 404 422 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 10 10 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 11 11 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 24 24 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 109 187 BY SIMILARITY.
FT VARIANT 16 16 P -> L.
FT VARIANT 22 22 /FTID-VAR_003446.
FT VARIANT 22 22 G -> S (IN DBSNP:1799920).
FT VARIANT 28 28 /FTID-VAR_011826.
FT VARIANT 184 184 I -> V (IN DBSNP:1799921).
FT VARIANT 184 184 P -> L (IN DBSNP:1800043).
FT VARIANT 220 220 /FTID-VAR_011828.
FT VARIANT 220 220 R -> L (IN DBSNP:1800044).
FT VARIANT 272 272 /FTID-VAR_011829.
FT VARIANT 273 273 G -> D.
FT VARIANT 273 273 G -> D (IN DBSNP:1800042).
FT CONFLICT 152 154 /FTID-VAR_011830.
FT CONFLICT 172 172 RAA -> PR (IN REF. 1).
FT CONFLICT 418 418 M -> I (IN REF. 1).
FT CONFLICT 418 418 K -> N (IN REF. 1).
SQ SEQUENCE 422 AA; 46106 MM; 762664FC62CFD8F CRC64;

Query Match 12.9%; Score 340.5; DB 1; Length 422;
Best Local Similarity 22.0%; Pred. No. 7.2e-13;
Matches 101; Conservative 92; Mismatches 190; Indels 77; Gaps 9;

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97324192; PubMed=9180361;
RA Miyamoto S., Taniguchi T., Suzuki F., Takita M., Kosaka N., Negoro E.,
RA Okuda T., Kosaka H., Mureta S., Nakamura S., Akagi Y., Oshita M.,
RA Watanabe Y., Muramatsu I.;
RT "Cloning, functional expression and tissue distribution of rabbit
alpha1a-adrenoceptor."
RL Life Sci. 60:2069-2074(1997).
CC - FUNCTION: THIS ALPHA-ADRENERGIC RECEPTOR MEDIATES ITS ACTION BY
CC ASSOCIATION WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM. ITS EFFECT IS MEDIATED BY G(Q)
CC AND G(11) PROTEINS.
CC - SUBCELLULAR LOCATION: Integral membrane protein.
CC - TISSUE SPECIFICITY: ABUNDANT IN LIVER, VAS DEFERENS, BRAIN, AND
CC AORTA, BUT NOT IN HEART.
CC - SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U81982; AAB6134.1; -.
DR GCRDP; GCR_1290; -.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm.1. 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G-PROTEIN_RECEPTOR_P1_1; 1.
DR PROSITE; PS00262; G-PROTEIN_RECEPTOR_P1_2; 1.
KW G-protein coupled receptor; Transmembrane; glycoprotein;
KW Multigene family; Phosphorylation; Lipoprotein; Palmitate.
FT DOMAIN 1 25 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 26 51 1 (POTENTIAL).
FT DOMAIN 52 63 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 64 89 2 (POTENTIAL).
FT DOMAIN 90 99 3 (POTENTIAL).
FT TRANSMEM 100 122 4 (POTENTIAL).
FT DOMAIN 123 143 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 144 168 5 (POTENTIAL).
FT DOMAIN 169 181 6 (POTENTIAL).
FT TRANSMEM 182 205 7 (POTENTIAL).
FT DOMAIN 206 272 8 (POTENTIAL).
FT TRANSMEM 273 297 9 (POTENTIAL).
FT DOMAIN 298 304 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 305 329 7 (POTENTIAL).
FT DOMAIN 330 466 CYTOPLASMIC (POTENTIAL).
FT LIPID 345 345 PALMITATE (POTENTIAL).
FT CARBOHYD 7 7 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 466 AA; 51365 MM; 9446D356B50FCD80 CRC64;

Query Match 12.9%; Score 340; DB 1; Length 466;
Best Local Similarity 22.0%; Pred. No. 8.5e-13;
Matches 102; Conservative 73; Mismatches 161; Indels 128; Gaps 9;

```

```

RESULT 3
ALIA_RABIT
ID ALIA_RABIT STANDARD; PRT; 466 AA.
AC 002824;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Alpha-1a adrenergic receptor (Alpha 1a-adrenoceptor) (Alpha-1C
DE adrenergic receptor).
GN ADRA1A OR ADRA1C.
OS Oryctolagus cuniculus (Rabbit).

```

```

FT DOMAIN 139 158 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 159 181 4 (POTENTIAL).
FT DOMAIN 182 203 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 204 226 5 (POTENTIAL).
FT DOMAIN 227 412 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 413 433 6 (POTENTIAL).
FT DOMAIN 434 447 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 448 467 7 (POTENTIAL).
FT DOMAIN 468 490 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 20 20 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 115 195 BY SIMILARITY.
SQ SEQUENCE 490 AA; 54937 MW; 2CDFDB5FA7D2298E CRC64;

Query Match 13.08; Score 343.5; DB 1; Length 490;
Best Local Similarity 23.7%; Pred. No. 5.7e-13;
Matches 116; Conservative 94; Mismatches 176; Indels 103; Gaps 17;

```

```

RX MEDLINE=87315369; PubMed=3041227;
RA Kobalka B.K., Friele T., Collins S., Yang-Feng T.L., Kobalka T.S.,
RA Franke U., Lefkowitz R.J., Caron M.G.;
RT "An intronless gene encoding a potential member of the family of
RL receptors coupled to guanine nucleotide regulatory proteins."
RL Nature 329:75-79(1987).
RN [2]
RP SEQUENCE FROM N.A.
RA Saltzman A.G., Morse B., Felder S.;
RL Submitted (FEB-1991) to the EMBL/GenBank/DBJ databases.
RN [3]
RP FUNCTION=
RX MEDLINE=88334700; PubMed=3138543;
RA Farqin A., Raymond J.R., Lobse M.L., Kobalka B.K., Caron M.G.,
RA Lefkowitz R.J.;
RT "The genomic clone G-21 which resembles a beta-adrenergic receptor
RL sequence encodes the 5-HT1A receptor."
RL Nature 335:358-360(1988).
RN [4]
RP VARIANT Asp-272.
RX MEDLINE=98425601; PubMed=9754630;
RA Kwanishi Y., Harada S., Tachikawa H., Okubo T., Shirashi H.;
RT "Novel mutations in the promoter and coding region of the human 5-HT1A
RL receptor gene and association analysis in schizophrenia."
RL Am. J. Med. Genet. 81:434-439(1998).
CC -1- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR 5-
CC HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC HORMONE THAT FUNCTIONS
CC AS A NEUROTRANSMITTER, A HORMONE, AND A MITOGEN. THE ACTIVITY OF
CC THIS RECEPTOR IS MEDIATED BY G PROTEINS THAT INHIBITS ADENYLATE
CC CYCLASE ACTIVITY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST TO THE OTHER 5HT-1 SUBTYPE RECEPTORS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M28269; AAA36440.1; -
DR EMBL; X13556; CAA31908.1; -
DR EMBL; X57829; CAA40962.1; -
DR EMBL; M83181; AAA66493.1; -
DR PIR; S07343; S07343.
DR PIR; S31438; S31438.
DR HSSP; P29274; 1MMH.
DR GCRDB; GCR_0087; -
DR GCRDB; GCR_0381; -
DR GCRDB; GCR_0624; -
DR MTM; 109760; -
DR InterPro; IPR00276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Multigene family; Polymorphism.
FT DOMAIN 1 36
FT TRANSMEM 63 73 1 (POTENTIAL).
FT DOMAIN 73 73 2 (POTENTIAL).
FT TRANSMEM 74 98 3 (POTENTIAL).
FT DOMAIN 99 109 4 (POTENTIAL).
FT TRANSMEM 110 132 5 (POTENTIAL).
FT DOMAIN 133 152 6 (POTENTIAL).
FT TRANSMEM 153 178 7 (POTENTIAL).
FT DOMAIN 179 191 8 (POTENTIAL).
FT TRANSMEM 192 217 9 (POTENTIAL).
FT DOMAIN 218 345 10 (POTENTIAL).
FT TRANSMEM 346 367 11 (POTENTIAL).
FT DOMAIN 368 378 12 (POTENTIAL).

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 16:41:57 : Search time 14 Seconds  
(without alignments)  
1404.967 Million cell updates/sec

Title: US-09-992-238-2

Perfect score: 2644  
Sequence: 1 MSTCTNSTRESNSSHTCMP.....GTEGTEGKTIVPSYDSTAFP 508

Scoring table: PROSUM62  
Gapop 10.0, Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	343.5	13.0	490	1	ACM4_CHICK
2	340.5	12.9	422	1	SH1A_HUMAN
3	340	12.9	466	1	ALIAA_RABIT
4	340	12.9	470	1	ALIAA_ORYLA
5	338.5	12.8	466	1	ALIAA_HUMAN
6	337.5	12.7	466	1	ALIAA_CAVPO
7	336	12.7	421	1	ALIAA_RAT
8	335.5	12.6	423	1	SH1A_MOUSE
9	333	12.6	466	1	SH1A_MOUSE
10	333	12.6	466	1	ALIAA_MOUSE
11	332.5	12.6	422	1	SH1A_RAT
12	325.5	12.3	466	1	SH1B_BOVIN
13	320	12.1	416	1	SH1B_FUGRU
14	319	12.1	509	1	5HT_LYMSR
15	317.5	12.0	562	1	ALIAA_MOUSE
16	314.5	11.9	572	1	ALIAA_HUMAN
17	314.5	11.6	561	1	ALIAA_RAT
18	304.5	11.5	484	1	ACM4_XENLA
19	304.5	11.5	476	1	OAR_DROME
20	300	11.3	579	1	ALIAA_RABIT
21	292	11.0	479	1	ACM4_HUMAN
22	290	11.0	466	1	ACM2_RAT
23	288	10.9	295	1	ALIAA_CANFA
24	288	10.9	488	1	YDWM_CABEL
25	287	10.9	466	1	ACM2_HUMAN
26	287	10.9	466	1	ACM2_MOUSE
27	283	10.7	515	1	ACM2_PIG
28	283	10.7	515	1	ALIAA_MESAU
29	283	10.7	515	1	ALIAA_RAT
30	283	10.7	519	1	ALIAA_HUMAN
31	282	10.7	479	1	ACM4_MOUSE
32	281.5	10.6	478	1	ACM4_RAT
33	280	10.6	466	1	ACM2_CHICK

34	279.5	10.6	476	1	GRE1_BALAM	O93126	balanus amp
35	277.5	10.5	531	1	ACM5_RAT	P08911	rattus norv
36	275.5	10.4	484	1	OAR1_LOCM1	O25321	locusta mig
37	275.5	10.4	532	1	ACM5_HUMAN	P08912	locusta mig
38	273.5	10.3	484	1	OAR2_LOCM1	O25322	locusta mig
39	273	10.3	443	1	D2DR_HUMAN	P14416	homo sapien
40	273	10.3	479	1	SH7_HUMAN	P34969	homo sapien
41	272.5	10.3	479	1	OAR_BOMMO	O17232	bombyx mori
42	272	10.3	446	1	SH7_CAVPO	P50407	cavia porce
43	272	10.3	436	1	SH71_DROME	P20905	drosophila
44	271.5	10.3	464	1	D2DR_MELGA	O73810	meleagris g
45	270	10.2	443	1	D2DR_CERAE	P52702	ceropithec

## ALIGNMENTS

RESULT 1	ID	ACM4_CHICK	STANDARD:	PRT:	490 AA.
AC	P17200	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Muscarinic acetylcholine receptor M4.				
GN	CHRM4.				
OS	Gallus gallus (Chicken).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;				
OC	Gallus.				
OX	NCBI_TaxID=9031;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90153912; PubMed=2154460;				
RA	Tieffe K.M., Goldman P.S., Nathanson N.M.;				
RT	"Cloning and functional analysis of a gene encoding a novel				
RT	muscarinic acetylcholine receptor expressed in chick heart and				
RT	brain.";				
RT	J. Biol. Chem. 265:2828-2834(1990).				
CC	- FUNCTION: THE MUSCARINIC ACETYLCHOLINE RECEPTOR MEDIATES VARIOUS				
CC	CELLULAR RESPONSES, INCLUDING INHIBITION OF ADENYLYL CYCLASE,				
CC	BREAKDOWN OF PHOSPHOLIPIDS & MODULATION OF POTASSIUM CHANNELS				
CC	THROUGH THE ACTION OF G PROTEINS. PRIMARY TRANSDUCING EFFECT IS				
CC	INHIBITION OF ADENYLYL CYCLASE. MAY COUPLE TO MULTIPLE FUNCTIONAL				
CC	RESPONSES IN CELL LINES.				
CC	- SUBCELLULAR LOCATION: Integral membrane protein.				
CC	- TISSUE SPECIFICITY: EXPRESSED IN HEART AND BRAIN.				
CC	- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.				
CC	- This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).				
CC	EMBL: J05218; AAA48563.1; -				
DR	PIR: A35546; A35546.				
DR	CCRD: GCR 0010: -				
DR	InterPro: IPR000276; GPCR_Rhodopsn.				
DR	Pfam: PF00001; 7tm_1; 1.				
DR	PRINTS: PR00237; GPCRHHODOPSN.				
DR	PROSITE: PS00237; G-PROTEIN_RECP_F1_1; 1.				
DR	PROSITE: PS50262; G-PROTEIN_RECP_F1_2; 1.				
KW	Postsynaptic membrane; Ionic channel; Glycoprotein; Transmembrane;				
KW	phosphorylation; Multigene family; G-protein coupled receptor.				
FT	DOMAIN 1 42				
FT	TRANSSEM 43 64				
FT	DOMAIN 65 78				
FT	TRANSSEM 79 99				
FT	DOMAIN 100 116				
FT	TRANSSEM 117 138				

Tue Oct 29 09:52:27 2002

us-09-992-238-2.rpr

Page 9

Db 339 -----KAAKTLAIYGVFVLCWPEFFFLPLGS-LFPQLKPSGSEVFVITMILGYFNS 389

Qy 448 CIIHRYVYGYMHTIKKEIQDMLK 470  
|::|:| :|:|:  
Db 390 CYNPLIYPCSSREFEKRAFLRLR 412

Search completed: October 28, 2002, 16:47:05  
Job time : 23 secs

THIS PAGE BLANK (USPTO)